

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**

Access DB# 99955

## RESEARCH REQUEST FORM

JUL 29 2003  
Scientific and Technical Information Center

Requester's Full Name: K.C. SRIVASTAVA Examiner #: 77964 Date: 07/29/2003  
 Art Unit: 1651 Phone Number 301-605-1196 Serial Number: 10/087,195  
 Mail Box and Bldg/Room Location: TIPSO 1 Results Format Preferred (circle): PAPER DISK E-MAIL  
ROOM # CM-11A12

If more than one search is submitted, please prioritize searches in order of need.

\*\*\*\*\*  
 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched.  
 Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: TREATMENT OF ALLERGIC RHINITISInventors (please provide full names): PATRICIA ANNE NUTALL &  
GUIDO CHRISTIAN PAESENEarliest Priority Filing Date: 09/01/1999

\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Please search for Protein sequence

AA 182 MS-HBPI = SEQ ID # 8

AA 172 FS-HBPI = SEQ ID # 6

AA 171 ES-HBPI = SEQ ID # 7

D. RET 6 = ?

TICK = RHIPKEPHALUS <sup>SINUS</sup> EVERTI  
<sup>SANGUINEUS</sup>And all terms ~~related~~ with ALLERGIC  
RHINITIS  
+ TREATMENT

Due 09/15/2003

+ above SEQ ID'S  
HISTACALIN PROTEIN" ECTOPARASITE (TICK  
OR - AMBLYOMMA SPP OR  
BLOOD FEEDING AMBLYOMMA AMERICANUM" SEASONAL OR PERENNIAL  
OR CHRONIC OR ACUTE  
+ ALLERGIC RHINITIS

## STAFF USE ONLY

Searcher: Hanby Type of Search: NA Sequence (#) STN \$182  
 Searcher Phone #: 301-605-1196 AA Sequence (#) 34 Dialog 02  
 Searcher Location: 8/1 Structure (#) Questel/Orbit 02  
 Date Searcher Picked Up: 8/1 Bibliographic Dr. Link 02  
 Date Completed: 8/1 Litigation Lexis/Nexis 02  
 Searcher Prep & Review Time: SEQ 20, STN 30 Fulltext Sequence Systems 02  
 Clerical Prep Time: 0 Patent Family WWW/Internet 02  
 Online Time: SEQ 20, STN 35 Other Other (specify) 02

**THIS PAGE BLANK (USPTO)**





# ***STIC Search Report***

## ***Biotech-Chem Library***

**STIC Database Tracking Number: 99955**

**TO: Kailash C Srivastava  
Location: cm-1/11a12/11b01  
Art Unit : 1651  
Monday, August 04, 2003**

**Case Serial Number: 10/087195**

**From : Susan Hanley  
Location: Biotech-Chem Library  
CM1 6B05  
Phone: 305-4053**

**susan.hanley@uspto.gov**

### **Search Notes**

**THIS PAGE BLANK (USPTO)**



# STIC SEARCH RESULTS FEEDBACK FORM

## Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher* or contact:

Mary Hale, Information Branch Supervisor  
308-4258, CM1-1E01

## Voluntary Results Feedback Form

➤ I am an examiner in Workgroup:  Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature  
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem/Library CM1 - Circ. Desk



**THIS PAGE BLANK (USPTO)**

# Registry Records for FSHBP1 & 2 & MSHBP1 & 2

D. RET 6

SRIVASTAVA 10/087,195

=> d sqide 1

L6 ANSWER 1 OF 3 REGISTRY COPYRIGHT 2003 ACS on STN **MSHBP1**  
 RN 329085-10-1 REGISTRY  
 CN 3: PN: WO0116164 TABLE: 1 unclaimed protein (9CI) (CA INDEX NAME)  
 OTHER NAMES:  
 CN 3: PN: WO0115719 TABLE: 1 unclaimed protein  
 FS PROTEIN SEQUENCE  
 SQL 182

## PATENT ANNOTATIONS (PNTE):

Sequence	Patent
Source	Reference
Not Given	WO2001015719
	unclaimed
	TABLE 1
	WO2001016164
	unclaimed
	TABLE 1

SEQ 1 NPTWANEAKL GSYQDAWKS QDQNKRYL AQATQTTDGV WGEEFTCVSV  
 51 TAEKIGKKKL NATILYKNKH LTDLKESHET ITWKAYDYT TENGIKYEYQ  
 101 GTRTQTFEDV FVFSYKNC D VIFVPKERGS DEGDYELWVS EDKIDKIPDC  
 151 CKFTMAYFAQ QKEKTVRNVY TDSSCKPAPA QN  
 MF Unspecified  
 CI MAN  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER, USPATFULL  
 2 REFERENCES IN FILE CA (1947 TO DATE)  
 2 REFERENCES IN FILE CAPLUS (1947 TO DATE)

=> d sqide 2

L6 ANSWER 2 OF 3 REGISTRY COPYRIGHT 2003 ACS on STN **FSHBP2**  
 RN 329085-09-8 REGISTRY  
 CN 2: PN: WO0116164 TABLE: 1 unclaimed sequence (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE  
 SQL 171

## PATENT ANNOTATIONS (PNTE):

Sequence	Patent
Source	Reference
Not Given	WO2001016164
	unclaimed
	TABLE 1

SEQ 1 NQPDWADEAA NGAHQDAWKS LKADVENVY MVKATYKNDP VWGNDFTCVG  
 51 VMANDVNEDE KSIQAEFLM NNADTNMQFA TEKVAVKMY GYNRENAFRY  
 101 ETEDGQVFTD VIAYSDDNCD VIYVPGTDGN EEGYELWTTD YDNIPANCLN  
 151 KFNEYAVGRE TRDVFTSACL E  
 MF Unspecified  
 CI MAN  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER, USPATFULL

**THIS PAGE BLANK (USPTO)**

1 REFERENCES IN FILE CA (1947 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1947 TO DATE)

=> d sqide 3

L6 ANSWER 3 OF 3 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 329085-08-7 REGISTRY  
CN 1: PN: WO0116164 TABLE: 1 unclaimed protein (9CI) (CA INDEX NAME)  
OTHER NAMES:  
CN 1: PN: WO0115719 TABLE: 1 unclaimed protein  
FS PROTEIN SEQUENCE  
SQL 172

FS-4BPI

PATENT ANNOTATIONS (PNTE):

Sequence	Patent
Source	Reference
Not Given	WO2001015719
	unclaimed
	TABLE 1
	WO2001016164
	unclaimed
	TABLE 1

SEQ 1 DKPVWADEAA NGEHQDAWKH LQKLVEENYD LIKATYKNDP VWGNDFTCVG  
51 TAAQNLNEDE KNVEAWFMFM NNADTVYQHT FEKATPDKMY GYNKENAITY  
101 QTEDGQVLTD VLAFSDDNCY VIYALGPDGS GAGYELWATD YTDVPASCLE  
151 KFNEYAAGLP VRDVYTSACL PE

MF Unspecified

CI MAN

SR CA

LC STN Files: CA, CAPLUS, TOXCENTER, USPATFULL  
2 REFERENCES IN FILE CA (1947 TO DATE)  
2 REFERENCES IN FILE CAPLUS (1947 TO DATE)

**THIS PAGE BLANK (USPTO)**



=> d sqide 128

D. RET6

L28 ANSWER 1 OF 1 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 200220-35-5 REGISTRY  
CN Protein D.RET6 (Dermacenter reticularis histamine-binding) (9CI) (CA  
INDEX NAME)  
FS PROTEIN SEQUENCE  
SQL 209

SEQ 1 MKMQVLLLT FVSAALATQA ETTSKAGEN PLWAHEELG KYQDAWSID  
51 QGVSVTYVLA KTTYENDTGS WGSQFKCLQV QEIERKEEDY TVTSVFTFRN  
101 ASSPIKYYNV TETVKAVFY GYKNIRNAIE YQVGGGLNIT DTLIFTDGEL  
151 CDVFYVPNAD QGCELWVKKS HYKHVPDYCT FVFNVCAKD RKTYDIFNEE  
201 CVYNGEPWL

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

MF Unspecified

CI MAN

SR CA

LC STN Files: CA, CAPLUS  
2 REFERENCES IN FILE CA (1947 TO DATE)  
2 REFERENCES IN FILE CAPLUS (1947 TO DATE)

**THIS PAGE BLANK (USPTO)**

This is a review on the HBP subject  
but there are no Registry #s associated  
w/ it

SRIVASTAVA 10/087,195

=> d ibib abs hitstr ind

L31 ANSWER 1 OF 1 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2000:799221 HCAPLUS  
DOCUMENT NUMBER: 134:55073  
TITLE: Vector-host interactions in disease transmission  
AUTHOR(S): Nuttall, P. A.; Paesen, G. C.; Lawrie, C. H.; Wang, H.  
CORPORATE SOURCE: CEH Institute of Virology and Environmental  
Microbiology, Oxford, OX1 3SR, UK  
SOURCE: Journal of Molecular Microbiology and Biotechnology  
(2000), 2(4), 381-386  
CODEN: JMMBFF; ISSN: 1464-1801  
PUBLISHER: Horizon Scientific Press  
DOCUMENT TYPE: Journal; General Review  
LANGUAGE: English

AB A review with 56 refs. Tick-borne spirochetes include borreliae that cause Lyme disease and relapsing fever in humans. They survive in a triangle of parasitic interactions between the spirochete and its vertebrate host, the spirochete and its tick vector, and the host and the tick. Until recently, the significance of vector-host interactions in the transmission of arthropod-borne disease agents has been overlooked. However, there is now compelling evidence that the pharmacol. activity of tick saliva can have a profound effect on pathogen transmission both from infected tick to uninfected host, and from infected host to uninfected tick. The salivary glands of ticks provide a pharmacopoeia of anti-inflammatory, anti-hemostatic and anti-immune mols. These include bioactive proteins that control histamine, bind Igs, and inhibit the alternative complement cascade. The effect of these mols. is to provide a privileged site at the tick-host interface in which borreliae and other tick-borne pathogens are sheltered from the normal innate and acquired host immune mechanisms that combat infections. Understanding the key events at the tick vector-host interface, that promote spirochete infection and transmission, will provide a better understanding of the epidemiol. and ecol. of these important human pathogens.

CC 15-0 (Immunochemistry)

Section cross-reference(s): 12

ST review immunomodulator tick Borrelia transmission

IT Proteins, specific or class

RL: BOC (Biological occurrence); BSU (Biological study, unclassified);  
BIOL (Biological study); OCCU (Occurrence)

(Ig-binding proteins; tick vector-host interaction and Borrelia  
transmission in relation to salivary expression of)

IT Complement

(activation; tick vector-host interaction and Borrelia transmission in  
relation to salivary expression of inhibitors of)

IT Proteins, specific or class

RL: BOC (Biological occurrence); BSU (Biological study, unclassified);  
BIOL (Biological study); OCCU (Occurrence)

(histacalins; tick vector-host interaction and Borrelia  
transmission in relation to salivary expression of)

IT Borrelia

Ixodes

Lyme disease

Saliva

(salivary immunomodulators in tick vector-host interaction and Borrelia  
transmission)

REFERENCE COUNT: 56 THERE ARE 56 CITED REFERENCES AVAILABLE FOR THIS  
RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

**THIS PAGE BLANK (USPTO)**

=> d que 132

L5 4 SEA FILE=REGISTRY ABB=ON PLU=ON (329085-08-7/BI OR 329085-09-8/BI OR 329085-10-1/BI OR 51-45-6/BI)  
L6 3 SEA FILE=REGISTRY ABB=ON PLU=ON L5 AND PROTEIN/FS  
L9 3 SEA FILE=HCAPLUS ABB=ON PLU=ON HISTACALIN?  
L10 2 SEA FILE=HCAPLUS ABB=ON PLU=ON L6  
L11 4 SEA FILE=HCAPLUS ABB=ON PLU=ON FS-HBP1  
L13 4 SEA FILE=HCAPLUS ABB=ON PLU=ON FS-HBP2  
L14 4 SEA FILE=HCAPLUS ABB=ON PLU=ON MS-HBP1  
L15 4 SEA FILE=HCAPLUS ABB=ON PLU=ON "D.RET6"  
L17 18 SEA FILE=HCAPLUS ABB=ON PLU=ON HISTAMINE BINDING PROTEIN  
L18 4 SEA FILE=HCAPLUS ABB=ON PLU=ON HISTAMINE BINDING PROTEIN-1  
L19 4 SEA FILE=HCAPLUS ABB=ON PLU=ON HISTAMINE BINDING PROTEIN-2  
L24 6 SEA FILE=HCAPLUS ABB=ON PLU=ON (L9 OR L10) OR L11 OR (L13 OR L14 OR L15) OR (L18 OR L19)

*citations  
w/ Reg #'s*

~~L32 14 SEA FILE=HCAPLUS ABB=ON PLU=ON L17 NOT L24~~

*← cites on HBP proteins  
but don't have the  
Registry #'s for  
D.RET6, SFHBP1, 2 or  
MS HBP1*

**THIS PAGE BLANK (USPTO)**

=&gt; d ibib abs hitstr 1-14 ind

L32 ANSWER 1 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2003:492892 HCAPLUS

TITLE: RNA interference in ticks: a study using  
**histamine binding protein**dsRNA in the female tick *Amblyomma americanum*AUTHOR(S): Aljamali, M. N.; Bior, A. D.; Sauer, J. R.; Essenberg,  
R. C.CORPORATE SOURCE: Department Biochemistry and Molecular Biology,  
Oklahoma State University, Stillwater, OK, 74078, USA

SOURCE: Insect Molecular Biology (2003), 12(3), 299-305

CODEN: IMBIE3; ISSN: 0962-1075

PUBLISHER: Blackwell Publishing Ltd.

DOCUMENT TYPE: Journal

LANGUAGE: English

AB RNA interference (RNAi), a gene silencing process, has been recently exploited to det. gene function by degrading specific mRNAs in several eukaryotic organisms. We constructed a double stranded RNA (dsRNA) from a previously cloned putative *Amblyomma americanum* **histamine binding protein** (HBP) to test the significance of using this methodol. in the assessment of the function and importance of gene products in ectoparasitic ticks. The female salivary glands incubated in vitro with HBP dsRNA had a significantly lower histamine binding ability. In addn., the injection of HBP dsRNA into the unfed females led both to a reduced histamine binding ability in the isolated salivary glands and to an aberrant tick feeding pattern or host response. Mol. data demonstrated less expression of the HBP mRNA in the RNAi group. Taken together, these results suggest that RNAi might be an important tool for assessing the significance of tick salivary gland secreted proteins modulating responses at the tick-host interface.

IT INDEXING IN PROGRESS

CC 3 (Biochemical Genetics)

IT *Amblyomma americanum*

Protein sequences

cDNA sequences

(rRNA interference in ticks, a study using **histamine binding protein** dsRNA in the female tick *Amblyomma americanum*)

IT 504653-46-7, GenBank AY246557

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL  
(Biological study)(nucleotide sequence; rRNA interference in ticks, a study using  
**histamine binding protein** dsRNA in the  
female tick *Amblyomma americanum*)REFERENCE COUNT: 30 THERE ARE 30 CITED REFERENCES AVAILABLE FOR THIS  
RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L32 ANSWER 2 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2002:385176 HCAPLUS

DOCUMENT NUMBER: 137:229499

TITLE: Comparison of differentially expressed genes in the  
salivary glands of male ticks, *Amblyomma americanum*  
and *Dermacentor andersoni*AUTHOR(S): Bior, Abdelaziz D.; Essenberg, Richard C.; Sauer, John  
R.CORPORATE SOURCE: 246 Noble Research Center, Department of Biochemistry  
and Molecular Biology, Oklahoma State University,  
Stillwater, OK, 74078-3035, USA

**THIS PAGE BLANK (USPTO)**



SOURCE: Insect Biochemistry and Molecular Biology (2002),  
32(6), 645-655  
CODEN: IBMBES; ISSN: 0965-1748  
PUBLISHER: Elsevier Science Ltd.  
DOCUMENT TYPE: Journal  
LANGUAGE: English

- AB Genes expressed differentially in the salivary glands of unfed and fed male ticks, *Amblyomma americanum* (L.), were identified, cloned and sequenced, and some were compared with those expressed in the salivary glands of *Dermacentor andersoni*. Total protein and RNA increased sixfold in the salivary glands of fed male *A. americanum*, while in fed male *D. andersoni* salivary glands, RNA increased approx. 3.5 times. Feeding *D. andersoni* in the presence of females increased total RNA by 25% over those fed in the absence of females. Complementary DNAs were synthesized from RNA obtained from unfed and fed ticks and amplified using RNA arbitrarily primed polymerase chain reaction (RAP-PCR) with three different primers in sep. reactions. Differential display showed clear banding differences between the fed and the unfed ticks in *A. americanum* and *D. andersoni*. Sixty-one cDNA fragments that appeared to be from differentially expressed genes in *A. americanum* were isolated, cloned and sequenced. Hybridization reactions with labeled cDNA probes confirmed the differential expression of many of the genes in unfed and fed ticks' salivary glands; however, many of the bands contained more than one fragment and some of the fragments isolated from apparently differential bands were not specific. Sequences for 28 of the cDNA fragments (150-600 nucleotides in length) demonstrated similarity to genes in the databases, but nine of these were similar to sequences of unknown function. Some of the gene fragments identified may be important to tick feeding or tick salivary gland physiol., including a histamine-binding protein, an org. ion transporter, an apoptosis inhibitor, a cathepsin-B-like cysteine protease, proteins involved in gene regulation and several proteins involved in protein synthesis. Cross-hybridization of identified cDNAs from *A. americanum* with cDNA probes synthesized from *D. andersoni* total RNA did not show significant similarity between the two species.
- CC 12-2 (Nonmammalian Biochemistry)  
Section cross-reference(s): 3
- ST gene expression salivary gland tick *Amblyomma* *Dermacentor* male feeding;  
sequence expressed sequence tag male feeding salivary gland tick
- IT *Amblyomma americanum*  
*Dermacentor andersoni*  
Salivary gland  
Sex  
(comparison of differentially expressed genes in the salivary glands of male ticks, *Amblyomma americanum* and *Dermacentor andersoni*)
- IT mRNA  
RL: BSU (Biological study, unclassified); BIOL (Biological study)  
(comparison of differentially expressed genes in the salivary glands of male ticks, *Amblyomma americanum* and *Dermacentor andersoni*)
- IT EST (expressed sequence tag)  
Gene, animal  
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)  
(comparison of differentially expressed genes in the salivary glands of male ticks, *Amblyomma americanum* and *Dermacentor andersoni*)
- IT cDNA sequences  
(for differentially expressed genes in the salivary glands of male ticks, *Amblyomma americanum* and *Dermacentor andersoni*)
- IT Blood  
(meal of; comparison of differentially expressed genes in the salivary

**THIS PAGE BLANK (USPTO)**

glands of male ticks, *Amblyomma americanum* and *Dermacentor andersoni*)  
 IT Proteins  
 RL: BSU (Biological study, unclassified); BIOL (Biological study)  
 (salivary gland; comparison of differentially expressed genes in the  
 salivary glands of male ticks, *Amblyomma americanum* and *Dermacentor  
 andersoni*)

IT 347782-65-4, GenBank BI273529	347782-66-5, GenBank BI273530
347782-67-6, GenBank BI273531	347782-68-7, GenBank BI273532
347782-69-8, GenBank BI273542	347782-70-1, GenBank BI273543
347782-71-2, GenBank BI273544	347782-72-3, GenBank BI273545
347782-73-4, GenBank BI273546	347782-74-5, GenBank BI273547
347782-75-6, GenBank BI273548	347782-76-7, GenBank BI273549
347782-77-8, GenBank BI273550	347782-78-9, GenBank BI273533
347782-79-0, GenBank BI273534	347782-80-3, GenBank BI273535
347782-81-4, GenBank BI273536	347782-82-5, GenBank BI273537
347782-83-6, GenBank BI273538	347782-84-7, GenBank BI273539
347782-85-8, GenBank BI273540	347782-86-9, GenBank BI273541
347782-87-0, GenBank BI273551	347782-88-1, GenBank BI273552
347782-89-2, GenBank BI273554	347782-90-5, GenBank BI273555
347782-91-6, GenBank BI273556	347782-92-7, GenBank BI273557
347782-93-8, GenBank BI273558	347782-94-9, GenBank BI273559
347782-95-0, GenBank BI273560	347782-96-1, GenBank BI273561
347782-97-2, GenBank BI273562	347782-98-3, GenBank BI273563
347782-99-4, GenBank BI273564	347783-00-0, GenBank BI273565
347783-01-1, GenBank BI273566	347783-02-2, GenBank BI273567
347783-03-3, GenBank BI273568	347783-04-4, GenBank BI273569
347783-05-5, GenBank BI273570	347783-06-6, GenBank BI273571
347783-07-7, GenBank BI273572	347783-08-8, GenBank BI273573
347783-09-9, GenBank BI273574	347783-10-2, GenBank BI273575
347783-11-3, GenBank BI273576	347783-12-4, GenBank BI273577
347783-13-5, GenBank BI273578	347783-14-6, GenBank BI273579
347783-15-7, GenBank BI273580	347783-16-8, GenBank BI273581
347783-17-9, GenBank BI273582	347783-18-0, GenBank BI273583
347783-19-1, GenBank BI273584	347783-20-4, GenBank BI273585
347783-21-5, GenBank BI273586	347783-22-6, GenBank BI273587
347783-23-7, GenBank BI273588	347783-24-8, GenBank BI273589
347783-25-9, GenBank BI273590	347783-26-0, GenBank BI273591
347783-27-1, GenBank BI273592	347783-28-2, GenBank BI273593
347783-29-3, GenBank BI273594	347783-30-6, GenBank BI273595
347783-31-7, GenBank BI273596	347783-32-8, GenBank BI273597
347783-33-9, GenBank BI273598	347783-34-0, GenBank BI273599
347783-35-1, GenBank BI273600	347783-36-2, GenBank BI273601
347783-37-3, GenBank BI273602	347783-38-4, GenBank BI273603
347783-39-5, GenBank BI273604	347783-40-8, GenBank BI273605
347783-41-9, GenBank BI273606	347783-42-0, GenBank BI273607
347783-43-1, GenBank BI273608	347783-44-2, GenBank BI273609
347783-45-3, GenBank BI273610	347783-46-4, GenBank BI273611
347783-47-5, GenBank BI273612	347783-48-6, GenBank BI273613
347783-49-7, GenBank BI273614	348411-30-3, GenBank BI275494
348411-31-4, GenBank BI275495	348411-32-5, GenBank BI275496
348411-33-6, GenBank BI275497	348411-34-7, GenBank BI275498
348411-35-8, GenBank BI275499	348411-36-9, GenBank BI275500
348411-37-0, GenBank BI275501	348411-38-1, GenBank BI275502
348411-39-2, GenBank BI275503	348411-40-5, GenBank BI275504
348411-41-6, GenBank BI275505	348411-42-7, GenBank BI275506
348411-43-8, GenBank BI275507	348411-44-9, GenBank BI275508
348411-45-0, GenBank BI275509	348411-46-1, GenBank BI275510
348411-47-2, GenBank BI275511	348411-48-3, GenBank BI275512
348411-49-4, GenBank BI275513	386326-31-4, GenBank BI273553

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL

**THIS PAGE BLANK (USPTO)**

## (Biological study)

(nucleotide sequence; comparison of differentially expressed genes in the salivary glands of male ticks, *Amblyomma americanum* and *Dermacentor andersoni*)

REFERENCE COUNT: 61 THERE ARE 61 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L32 ANSWER 3 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2002:231098 HCAPLUS

DOCUMENT NUMBER: 136:383078

TITLE: A high affinity serotonin- and histamine-binding lipocalin from tick saliva

AUTHOR(S): Sangamnatdej, S.; Paesen, G. C.; Slovak, M.; Nuttall, P. A.

CORPORATE SOURCE: CEH Oxford, Oxford, OX1 3SR, UK

SOURCE: Insect Molecular Biology (2002), 11(1), 79-86

CODEN: IMBIE3; ISSN: 0962-1075

PUBLISHER: Blackwell Publishing Ltd.

DOCUMENT TYPE: Journal

LANGUAGE: English

AB To overcome the inflammatory response in its host, the cattle-feeding, brown ear tick secretes **histamine-binding proteins** into the feeding site. These proteins are .beta.-barrels with 2 internal binding sites: a high-affinity (H) site for histamine and a site (L) for which the natural ligand is unknown. Here we report a related protein (SHBP), secreted by a rodent- and cattle-feeding tick, that traps both histamine and serotonin. The histamine-binding H site is well conserved in SHBP, whereas residue changes in the L-like site are consistent with binding of the bulkier serotonin mol. As histamine is a key inflammatory mediator in cattle, while serotonin takes on this role in rodents, the diversification of these tick proteins may reflect host adaptation.

CC 12-1 (Nonmammalian Biochemistry)

Section cross-reference(s): 3, 6

ST Dermacentor saliva serotonin histamine binding lipocalin sequence

IT Dermacentor reticulatus

Feeding

Protein sequences

Saliva

cDNA sequences

(high affinity serotonin- and histamine-binding lipocalin sequence and identification from tick saliva)

IT Proteins

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL

(Biological study)

(lipocalin; high affinity serotonin- and histamine-binding lipocalin sequence and identification from tick saliva)

IT 426310-26-1, Lipocalin (Dermacentor reticulatus)

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL

(Biological study)

(amino acid sequence; high affinity serotonin- and histamine-binding lipocalin sequence and identification from tick saliva)

IT 50-67-9, Serotonin, biological studies 51-45-6, Histamine, biological studies

RL: BSU (Biological study, unclassified); BIOL (Biological study)

(high affinity serotonin- and histamine-binding lipocalin sequence and identification from tick saliva)

IT 381905-30-2, GenBank AF217101

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL

(Biological study)

**THIS PAGE BLANK (USPTO)**

These are all of the cites for the Registry #15 for FSHBP1, 2, MSHBP1 & D.RET6

FSHBP1

SRIVASTAVA 10/087,195

FSHBP1

FSHBP-2

=> d que 130

L5 4 SEA FILE=REGISTRY ABB=ON PLU=ON (329085-08-7/BI OR 329085-09-8/BI OR ~~329085-10-1/BI OR 51-45-6/BI~~)

L6 3 SEA FILE=REGISTRY ABB=ON PLU=ON L5 AND PROTEIN/FS

L9 3 SEA FILE=HCAPLUS ABB=ON PLU=ON HISTACALIN? 3 cites for Histacalin

L10 2 SEA FILE=HCAPLUS ABB=ON PLU=ON L6 2 cites for FSHBP1, 2 & MSHBP1 from

L11 4 SEA FILE=HCAPLUS ABB=ON PLU=ON FS-HBP1

L13 4 SEA FILE=HCAPLUS ABB=ON PLU=ON FS-HBP2

L14 4 SEA FILE=HCAPLUS ABB=ON PLU=ON MS-HBP1 } term searching

L15 4 SEA FILE=HCAPLUS ABB=ON PLU=ON "D.RET6" }

L18 4 SEA FILE=HCAPLUS ABB=ON PLU=ON HISTAMINE BINDING PROTEIN-1 }

L19 4 SEA FILE=HCAPLUS ABB=ON PLU=ON HISTAMINE BINDING PROTEIN-2 }

L24 6 SEA FILE=HCAPLUS ABB=ON PLU=ON (L9 OR L10) OR L11 OR (L13 OR L14 OR L15) OR (L18 OR L19) 6 cites total

L25 31 SEA FILE=REGISTRY ABB=ON PLU=ON (51-45-6/BI OR 200220-28-6/BI OR 200220-29-7/BI OR 200220-30-0/BI OR 200220-32-2/BI OR 200220-33-3/BI OR 200220-34-4/BI OR 200220-31-1/BI OR 200220-35-5/BI OR 329085-08-7/BI OR 329085-10-1/BI OR 226219-91-6/BI OR 226890-87-5/BI OR 226890-96-6/BI OR 226890-97-7/BI OR 226890-98-8/BI OR 226891-00-5/BI OR 226891-16-3/BI OR 226891-31-2/BI OR 226891-34-5/BI OR 226891-36-7/BI OR 226891-37-8/BI OR 226891-38-9/BI OR 226891-39-0/BI OR 226891-40-3/BI OR 226927-07-7/BI OR 329085-09-8/BI OR 329085-11-2/BI OR 50-67-9/BI OR 501-75-7/BI OR 644-42-8/BI) } 31 compounds in the combined L24 cites

L26 16 SEA FILE=REGISTRY ABB=ON PLU=ON L25 AND PROTEIN/FS 16 are proteins

L27 5 SEA FILE=HCAPLUS ABB=ON PLU=ON L24 AND L26 5 cites

L28 1 SEA FILE=REGISTRY ABB=ON PLU=ON L26 AND SQL=209 ← D.RET6

L29 2 SEA FILE=HCAPLUS ABB=ON PLU=ON L28 2 cites for

L30 5 SEA FILE=HCAPLUS ABB=ON PLU=ON L29 FOR L27 5 cites total

FS = file segment

**THIS PAGE BLANK (USPTO)**



=&gt; d ibib abs hitstr ind 1-5 130

L30 ANSWER 1 OF 5 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2001:168020 HCAPLUS

DOCUMENT NUMBER: 134:217189

TITLE: Treatment of allergic rhinitis with proteins from ticks

INVENTOR(S): Nuttall, Patricia Anne; Paesen, Guido Christiaan

PATENT ASSIGNEE(S): Evlutec Limited, UK

SOURCE: PCT Int. Appl., 19 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001016164	A2	20010308	WO 2000-GB3287	20000824
WO 2001016164	A3	20010503		

W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM

RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

BR 2000013655	A	20020507	BR 2000-13655	20000824
---------------	---	----------	---------------	----------

EP 1207899	A2	20020529	EP 2000-954788	20000824
------------	----	----------	----------------	----------

R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO, MK, CY, AL

JP 2003508410	T2	20030304	JP 2001-519725	20000824
---------------	----	----------	----------------	----------

US 2002193306	A1	20021219	US 2002-87195	20020301
---------------	----	----------	---------------	----------

PRIORITY APPLN. INFO.:

GB 1999-20673 A 19990901

WO 2000-GB3287 W 20000824

AB The invention relates to the discovery that various proteins isolated from ticks are effective in the treatment and prevention of allergic rhinitis. These proteins may most suitably be applied to an effected area and are thus effective to treat this condition and to ameliorate its symptoms. Human subjects were challenged with histamine and then were treated with histamine-binding protein, MS-HBP1.

IT 329085-08-7

RL: PRP (Properties)

(unclaimed protein sequence; treatment of allergic rhinitis with proteins from ticks)

RN 329085-08-7 HCAPLUS

CN 1: PN: WO0116164 TABLE: 1 unclaimed protein (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

IT 329085-09-8 329085-10-1

RL: PRP (Properties)

(unclaimed sequence; treatment of allergic rhinitis with proteins from ticks)

RN 329085-09-8 HCAPLUS

CN 2: PN: WO0116164 TABLE: 1 unclaimed sequence (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

**THIS PAGE BLANK (USPTO)**

RN 329085-10-1 HCAPLUS

CN 3: PN: W00116164 TABLE: 1 unclaimed protein (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

IC ICM C07K014-00

CC 1-7 (Pharmacology)

Section cross-reference(s): 12, 15

ST allergic rhinitis treatment protein tick; histamine binding protein MSHBP1  
treatment allergic rhinitis

IT Proteins, specific or class

RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)

(D.RET6; treatment of allergic rhinitis with proteins from ticks)

IT Proteins, specific or class

RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)

(FS-HBP1 (female-specific histamine-binding protein 1); treatment of allergic rhinitis with proteins from ticks)

IT Proteins, specific or class

RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)

(FS-HBP2 (female-specific histamine-binding protein 2); treatment of allergic rhinitis with proteins from ticks)

IT Proteins, specific or class

RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)

(MS-HBP1 (male-specific histamine-binding protein 1); treatment of allergic rhinitis with proteins from ticks)

IT Nose

(allergic rhinitis; treatment of allergic rhinitis with proteins from ticks)

IT Nervous system stimulants

(antisedatives, medicament also contg.; treatment of allergic rhinitis with proteins from ticks)

IT Parasite

(ecto-, histacalin of; treatment of allergic rhinitis with proteins from ticks)

IT Proteins, specific or class

RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)

(histacalins; treatment of allergic rhinitis with proteins from ticks)

IT Antihistamines

(medicament also contg.; treatment of allergic rhinitis with proteins from ticks)

IT Proteins, general, biological studies

RL: THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
(of ticks; treatment of allergic rhinitis with proteins from ticks)

IT Allergy inhibitors

Drug delivery systems

Hay fever

**THIS PAGE BLANK (USPTO)**

## Mite and Tick

(treatment of allergic rhinitis with proteins from ticks)

IT 51-45-6, Histamine, biological studies

RL: ADV (Adverse effect, including toxicity); BPR (Biological process);  
BSU (Biological study, unclassified); BIOL (Biological study); PROC  
(Process)(proteins binding to; treatment of allergic rhinitis with proteins from  
ticks)

IT 329085-08-7

RL: PRP (Properties)

(unclaimed protein sequence; treatment of allergic rhinitis with  
proteins from ticks)

IT 329085-09-8 329085-10-1

RL: PRP (Properties)

(unclaimed sequence; treatment of allergic rhinitis with proteins from  
ticks)

L30 ANSWER 2 OF 5 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2001:167826 HCAPLUS

DOCUMENT NUMBER: 134:217188

TITLE: use of **histacalin** protein for treatment or  
prevention of conjunctivitis

INVENTOR(S): Nuttall, Patricia Anne; Paesen, Guido Christiaan

PATENT ASSIGNEE(S): Evolutech Limited, UK

SOURCE: PCT Int. Appl., 19 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

## PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001015719	A2	20010308	WO 2000-GB3282	20000824
WO 2001015719	A3	20010510		

W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN,  
CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,  
HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT,  
LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU,  
SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN,  
YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM

RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY,  
DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ,  
CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

BR 2000013665 A 20020514 BR 2000-13665 20000824

EP 1207898 A2 20020529 EP 2000-954784 20000824

R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT,  
IE, SI, LT, LV, FI, RO, MK, CY, AL

JP 2003508445 T2 20030304 JP 2001-519931 20000824

US 2002151499 A1 20021017 US 2002-85572 20020227

PRIORITY APPLN. INFO.: GB 1999-20674 A 19990901

WO 2000-GB3282 W 20000824

AB Various **histacalin** proteins isolated from ticks are effective in  
the treatment of conjunctivitis. These proteins may most suitably be  
applied topically to an affected area and are effective to ameliorate the  
symptoms of this condition.

IT 329085-08-7 329085-10-1 329085-11-2

RL: PRP (Properties)

(unclaimed sequence; use of **histacalin** protein for treatment  
or prevention of conjunctivitis)

**THIS PAGE BLANK (USPTO)**

RN 329085-08-7 HCAPLUS  
 CN 1: PN: W00116164 TABLE: 1 unclaimed protein (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 329085-10-1 HCAPLUS  
 CN 3: PN: W00116164 TABLE: 1 unclaimed protein (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 329085-11-2 HCAPLUS  
 CN 2: PN: W00115719 TABLE: 1 unclaimed sequence (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

IC ICM A61K038-00  
 CC 1-7 (Pharmacology)  
 Section cross-reference(s): 63  
 ST **histacalin** antihistamine antiallergic antiinflammatory conjunctivitis  
 IT Proteins, specific or class  
 RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
 (D.RET6; **histacalin** protein for treatment or prevention conjunctivitis)  
 IT Proteins, specific or class  
 RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
 (FS-HBP1; **histacalin** protein for treatment or prevention conjunctivitis)  
 IT Proteins, specific or class  
 RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
 (FS-HBP2; **histacalin** protein for treatment or prevention conjunctivitis)  
 IT Proteins, specific or class  
 RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
 (MS-HBP1; **histacalin** protein for treatment or prevention conjunctivitis)  
 IT Eye, disease  
 (allergic conjunctivitis; **histacalin** protein for treatment or prevention conjunctivitis)  
 IT Eye, disease  
 (conjunctivitis; **histacalin** protein for treatment or prevention conjunctivitis)  
 IT Mast cell  
 (degranulation; **histacalin** protein for treatment or prevention conjunctivitis)  
 IT Parasite  
 (ecto-, blood-feeding; **histacalin** protein for treatment or prevention conjunctivitis)  
 IT Allergy inhibitors  
 Anti-inflammatory agents  
 Antihistamines  
 Drug delivery systems  
 Mite and Tick  
 Protein sequences  
 (**histacalin** protein for treatment or prevention)

**THIS PAGE BLANK (USPTO)**



IT conjunctivitis)  
 IT Proteins, specific or class  
 RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
 (histacalin; histacalin protein for treatment or prevention conjunctivitis)  
 IT Cell degranulation  
 (mast cell; histacalin protein for treatment or prevention conjunctivitis)  
 IT Mental activity  
 (sedation, antisedatives; histacalin protein for treatment or prevention conjunctivitis)  
 IT Drug delivery systems  
 (solns., ophthalmic; histacalin protein for treatment or prevention conjunctivitis)  
 IT 329085-08-7 329085-10-1 329085-11-2  
 RL: PRP (Properties)  
 (unclaimed sequence; use of histacalin protein for treatment or prevention of conjunctivitis)

L30 ANSWER 3 OF 5 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1999:374997 HCAPLUS

DOCUMENT NUMBER: 131:154999

TITLE: Tick histamine-binding proteins: isolation, cloning, and three-dimensional structure

AUTHOR(S): Paesen, G. C.; Adams, P. L.; Harlos, K.; Nuttall, P. A.; Stuart, D. I.

CORPORATE SOURCE: Institute of Virology and Environmental Microbiology, Natural Environment Research Council, Oxford, OX1 3SR, UK

SOURCE: Molecular Cell (1999), 3(5), 661-671

CODEN: MOCEFL; ISSN: 1097-2765

PUBLISHER: Cell Press

DOCUMENT TYPE: Journal

LANGUAGE: English

AB High-affinity histamine-binding proteins (HBPs) were discovered in the saliva of Rhipicephalus appendiculatus ticks. Their ability to outcompete histamine receptors indicates that they suppress inflammation during blood feeding. The crystal structure of a histamine-bound HBP, detd. at 1.25 Å. resoln., reveals a lipocalin fold novel in contg. two binding sites for the same ligand. The sites are orthogonally arranged and highly rigid and form an internal surface of unusual polar character that complements the physicochem. properties of histamine. As sol. receptors of histamine, HBPs offer a new strategy for controlling histamine-based diseases.

IT 200220-32-2 200220-33-3 200220-34-4

RL: PRP (Properties)

(amino acid sequence; isolation, cloning, mol. characterization and three-dimensional structure of sex-specific tick histamine-binding proteins)

RN 200220-32-2 HCAPLUS

CN Protein FS-HBP1 (Rhipicephalus appendiculatus female-specific histamine-binding) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 200220-33-3 HCAPLUS

CN Protein FS-HBP2 (Rhipicephalus appendiculatus female-specific histamine-binding) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

**THIS PAGE BLANK (USPTO)**

RN 200220-34-4 HCAPLUS

CN Protein MS-HBP1 (Rhipicephalus appendiculatus male-specific histamine-binding) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

CC 6-3 (General Biochemistry)

Section cross-reference(s): 3, 12

ST Rhipicephalus histamine binding protein structure; cDNA sequence tick histamine binding protein

IT Proteins, specific or class

RL: BAC (Biological activity or effector, except adverse); BOC (Biological occurrence); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); OCCU (Occurrence)

(HBP (histamine-binding proteins); isolation, cloning, mol.

characterization and three-dimensional structure of sex-specific tick histamine-binding proteins)

IT Proteins, specific or class

RL: BAC (Biological activity or effector, except adverse); BOC (Biological occurrence); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); OCCU (Occurrence)

(HBP-1 (histamine-binding protein

1); isolation, cloning, mol. characterization and three-dimensional structure of sex-specific tick histamine-binding proteins)

IT Proteins, specific or class

RL: BAC (Biological activity or effector, except adverse); BOC (Biological occurrence); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); OCCU (Occurrence)

(HBP-2 (histamine-binding protein

2); isolation, cloning, mol. characterization and three-dimensional structure of sex-specific tick histamine-binding proteins)

IT Crystal structure

Protein sequences

Rhipicephalus appendiculatus

cDNA sequences

(isolation, cloning, mol. characterization and three-dimensional structure of sex-specific tick histamine-binding proteins)

IT Tertiary structure

(protein; isolation, cloning, mol. characterization and three-dimensional structure of sex-specific tick histamine-binding proteins)

IT 200220-32-2 200220-33-3 200220-34-4

RL: PRP (Properties)

(amino acid sequence; isolation, cloning, mol. characterization and three-dimensional structure of sex-specific tick histamine-binding proteins)

IT 51-45-6, Histamine, biological studies

RL: BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PROC (Process)

(isolation, cloning, mol. characterization and three-dimensional structure of sex-specific tick histamine-binding proteins)

IT 200220-28-6, GenBank U96080 200220-29-7, GenBank U96081 200220-30-0, GenBank U96082

RL: PRP (Properties)

(nucleotide sequence; isolation, cloning, mol. characterization and three-dimensional structure of sex-specific tick histamine-binding proteins)

REFERENCE COUNT:

48

THERE ARE 48 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

**THIS PAGE BLANK (USPTO)**

L30 ANSWER 4 OF 5 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1999:359659 HCAPLUS  
 DOCUMENT NUMBER: 131:28315  
 TITLE: Cloning and functions of vasoactive amine-binding proteins from ticks  
 INVENTOR(S): Nuttall, Patricia Ann; Paesen, Guido Christian  
 PATENT ASSIGNEE(S): Oxford Vacs Ltd., UK  
 SOURCE: PCT Int. Appl., 84 pp.  
 CODEN: PIXXD2  
 DOCUMENT TYPE: Patent  
 LANGUAGE: English  
 FAMILY ACC. NUM. COUNT: 1  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9927104	A1	19990603	WO 1998-GB3530	19981126
W: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM RW: GH, GM, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
CA 2309809	AA	19990603	CA 1998-2309809	19981126
AU 9912511	A1	19990615	AU 1999-12511	19981126
EP 1034273	A1	20000913	EP 1998-955786	19981126
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI, RO				
BR 9815056	A	20001003	BR 1998-15056	19981126
JP 2002508927	T2	20020326	JP 2000-522246	19981126
NZ 504753	A	20021126	NZ 1998-504753	19981126
PRIORITY APPLN. INFO.:			GB 1997-25046	A 19971126
			GB 1998-13917	A 19980626
			WO 1998-GB3530	W 19981126

AB The present invention relates to histamine and serotonin binding mols. that possess a binding site with the precise mol. configuration that is necessary to confer on the mol. a high affinity for histamine. The invention includes proteins, peptides and chem. compds. that possess this mol. configuration and that are thus able to bind to histamine with high affinity. These mols. may be used in the regulation of the action of histamine or serotonin, the detection and quantification of histamine or serotonin and in the treatment of various diseases and allergies. The mols. may also be used as components of vaccines directed against blood-sucking ectoparasites. Vasoactive amine binding proteins (VABPs) are provided that specifically bind to vasoactive amines with a dissocn. const. of  $<10^{-7}$  M and which belong to the same protein family as MS-HBP1, FS-HBP1, FS-HBP2 and D.RET6. Thus, 11 VASPs were isolated, and their cDNAs cloned and sequenced, from ticks: FS-HBP1 (female-specific histamine-binding protein 1), FS-HBP2 (female-specific histamine-binding protein 2), MS-HBP1 (male-specific histamine-binding protein 1), and Ra-Res from Rhipicephalus appendiculatus; D.RET6 from Dermacentor reticularis; Av-HBP from Amblyomma variegatum; and 5 related Ih/Bm-HBP proteins from a mixed Ixodes hexagonus/Boophilus microplus cDNA expression

**THIS PAGE BLANK (USPTO)**

library. These VASPs possess similar amino acid sequences and predicted secondary structures. The VASPs bind histamine in mammals, and can be used as anti-inflammatory agents to regulate histamine action and to control its pathol. effects. The crystal structure of FS-HBP2 to 2.24 Å. resoln. was used to design a synthetic cyclic octapeptide (-Ala-Glu-Ala-Phe-Ala-Glu-Ala-Trp-) with histamine binding activity.

IT 200220-32-2 200220-33-3 200220-34-4  
200220-35-5 226891-31-2 226891-34-5  
226891-36-7 226891-37-8 226891-38-9  
226891-39-0 226891-40-3

RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses) (amino acid sequence; cloning and functions of vasoactive amine-binding proteins from ticks)

RN 200220-32-2 HCAPLUS

CN Protein FS-HBP1 (*Rhipicephalus appendiculatus* female-specific histamine-binding) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 200220-33-3 HCAPLUS

CN Protein FS-HBP2 (*Rhipicephalus appendiculatus* female-specific histamine-binding) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 200220-34-4 HCAPLUS

CN Protein MS-HBP1 (*Rhipicephalus appendiculatus* male-specific histamine-binding) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 200220-35-5 HCAPLUS

CN Protein D.RET6 (*Dermacenter reticularis* histamine-binding) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 226891-31-2 HCAPLUS

CN Histamine-binding protein Ra-Res (*Rhipicephalus appendiculatus*) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 226891-34-5 HCAPLUS

CN Histamine-binding protein Av-HBP (*Amblyomma variegatum*) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 226891-36-7 HCAPLUS

CN Histamine-binding protein Ih/Bm-HBP1 (*Boophilus microplus*/Ixodes hexagonus) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 226891-37-8 HCAPLUS

CN Histamine-binding protein Ih/Bm-HBP2 (*Boophilus microplus*/Ixodes hexagonus) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 226891-38-9 HCAPLUS

CN Histamine-binding protein Ih/Bm-HBP3 (*Boophilus microplus*/Ixodes hexagonus) (9CI) (CA INDEX NAME)

13 PAGE BLANK (USPTO)



\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 226891-39-0 HCAPLUS

CN Histamine-binding protein Ih/Bm-HBP4 (Boophilus microplus/Ixodes hexagonus) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 226891-40-3 HCAPLUS

CN Histamine-binding protein Ih/Bm-HBP5 (Boophilus microplus/Ixodes hexagonus) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

IT 226219-91-6

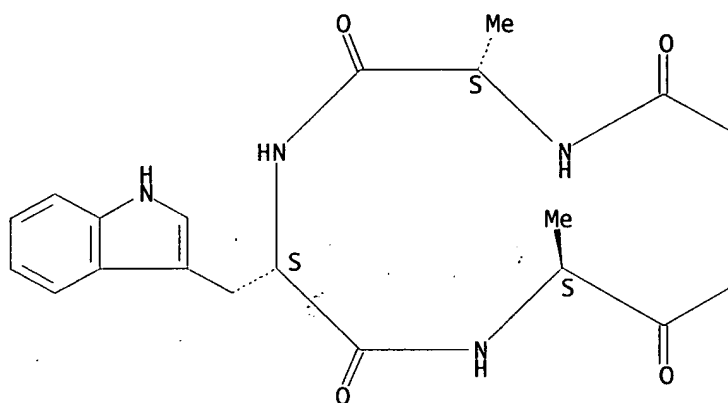
RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)  
(cloning and functions of vasoactive amine-binding proteins from ticks)

RN 226219-91-6 HCAPLUS

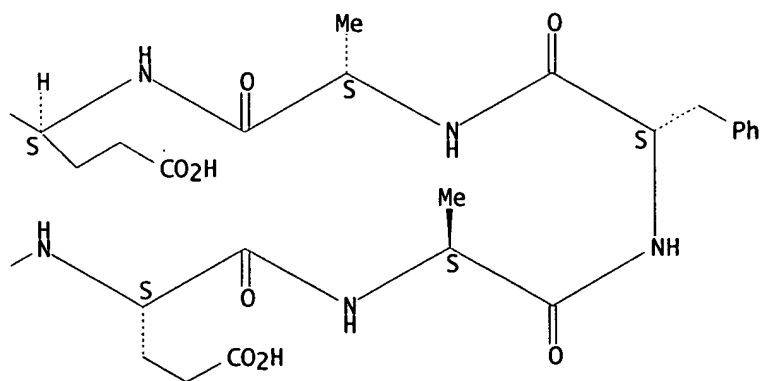
CN Cyclo(L-alanyl-L-.alpha.-glutamyl-L-alanyl-L-phenylalanyl-L-alanyl-L-.alpha.-glutamyl-L-alanyl-L-tryptophyl) (9CI) (CA INDEX NAME)

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



**THIS PAGE BLANK (USPTO)**

- IC ICM C12N015-21
- ICS C07K014-435; A61K038-17; A61K031-40; A61K031-19; A61K031-66;  
A61K031-35; G01N033-68; A23L001-015; C12N005-10; C12N001-21;  
A01K067-027; C12N015-00
- CC 2-8 (Mammalian Hormones)  
Section cross-reference(s): 1, 3, 6, 12, 75
- ST vasoactive amine binding protein tick; histamine binding protein tick;  
Rhipicephalus vasoactive amine binding protein; Dermacenter vasoactive  
amine binding protein; Amblyomma vasoactive amine binding protein;  
Boophilus vasoactive amine binding protein; Ixodes vasoactive amine  
binding protein; sequence vasoactive binding protein cDNA tick;  
inflammation inhibitor VASP protein tick
- IT Proteins, specific or class  
RL: BAC (Biological activity or effector, except adverse); BPR (Biological  
process); BSU (Biological study, unclassified); PRP (Properties); THU  
(Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)  
(D.RET6; cloning and functions of vasoactive  
amine-binding proteins from ticks)
- IT Proteins, specific or class  
RL: BAC (Biological activity or effector, except adverse); BPR (Biological  
process); BSU (Biological study, unclassified); PRP (Properties); THU  
(Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)  
(FS-HBP1 (female-specific histamine-  
binding protein 1); cloning and functions  
of vasoactive amine-binding proteins from ticks)
- IT Proteins, specific or class  
RL: BAC (Biological activity or effector, except adverse); BPR (Biological  
process); BSU (Biological study, unclassified); PRP (Properties); THU  
(Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)  
(FS-HBP2 (female-specific histamine-  
binding protein 2); cloning and functions  
of vasoactive amine-binding proteins from ticks)
- IT Proteins, specific or class  
RL: BAC (Biological activity or effector, except adverse); BPR (Biological  
process); BSU (Biological study, unclassified); PRP (Properties); THU  
(Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)  
(Ih/Bm-HBP (Ixodes hexagonus/Boophilus microplus histamine-binding  
protein); cloning and functions of vasoactive amine-binding proteins  
from ticks)
- IT Proteins, specific or class  
RL: BAC (Biological activity or effector, except adverse); BPR (Biological  
process); BSU (Biological study, unclassified); PRP (Properties); THU  
(Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)  
(MS-HBP1 (male-specific histamine-  
binding protein 1); cloning and functions  
of vasoactive amine-binding proteins from ticks)
- IT Proteins, specific or class  
RL: BAC (Biological activity or effector, except adverse); BPR (Biological  
process); BSU (Biological study, unclassified); PRP (Properties); THU  
(Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)  
(VASP (vasoactive amine-binding protein); cloning and functions of  
vasoactive amine-binding proteins from ticks)
- IT Amines, biological studies  
RL: BPR (Biological process); BSU (Biological study, unclassified); BIOL  
(Biological study); PROC (Process)  
(biogenic, vasoactive; cloning and functions of vasoactive  
amine-binding proteins from ticks)
- IT 5-HT antagonists  
Amblyomma variegatum

**THIS PAGE BLANK (USPTO)**

- Anti-inflammatory agents
- Antihistamines
- Boophilus microplus
- Dermacentor reticulatus
- Ixodes hexagonus
- Mite and Tick
- Molecular cloning
- Rhipicephalus appendiculatus
- Scorpion
- Snake
- Spider
- Transformation, genetic
- Vaccines
  - (cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Fusion proteins (chimeric proteins)
  - RL: BPN (Biosynthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)
  - (cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Peptides, biological studies
  - RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)
  - (cyclic; cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Parasite
  - (ecto-; cloning and functions of vasoactive amine-binding proteins from ticks)
- IT cDNA sequences
  - (for vasoactive amine-binding proteins from ticks)
- IT Food analysis
- Plant analysis
  - (histamine detn. in; cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Diagnosis
  - (mol.; cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Crystal structure
  - (of histamine-binding protein FS-HBP2; cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Protein sequences
  - (of vasoactive amine-binding proteins from ticks)
- IT Conformation
  - (protein, of histamine-binding protein FS-HBP2; cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Body fluid
  - (quantification of histamine levels in body fluids; cloning and functions of vasoactive amine-binding proteins from ticks)
- IT 200220-32-2 200220-33-3 200220-34-4  
 200220-35-5 226891-31-2 226891-34-5  
 226891-36-7 226891-37-8 226891-38-9  
 226891-39-0 226891-40-3
  - RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)
  - (amino acid sequence; cloning and functions of vasoactive amine-binding proteins from ticks)
- IT 50-67-9, Serotonin, biological studies 51-45-6, Histamine, biological studies 501-75-7, 1-Methylhistamine 644-42-8, 3-Methylhistamine
  - RL: ANT (Analyte); BPR (Biological process); BSU (Biological study, unclassified); ANST (Analytical study); BIOL (Biological study); PROC

**THIS PAGE BLANK (USPTO)**

(Process)

(cloning and functions of vasoactive amine-binding proteins from ticks)

IT 226219-91-6

RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)

(cloning and functions of vasoactive amine-binding proteins from ticks)

IT 200220-28-6P 200220-29-7P 200220-30-0P 200220-31-1P 226890-87-5P  
226890-96-6P 226890-97-7P 226890-98-8P 226891-00-5P 226891-16-3P  
226927-07-7P

RL: BPN (Biosynthetic preparation); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)

(nucleotide sequence; cloning and functions of vasoactive amine-binding proteins from ticks)

REFERENCE COUNT: 5 THERE ARE 5 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L30 ANSWER 5 OF 5 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1997:776255 HCAPLUS

DOCUMENT NUMBER: 128:57765

TITLE: Cloning and functions of vasoactive amine-binding proteins from ticks

INVENTOR(S): Paesen, Guido Christian; Nuttall, Patricia Ann

PATENT ASSIGNEE(S): Oxford Vacs Ltd., UK; Paesen, Guido Christian; Nuttall, Patricia Ann

SOURCE: PCT Int. Appl., 44 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9744451	A2	19971127	WO 1997-GB1372	19970519
WO 9744451	A3	19980219		

W: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM

RW: GH, KE, LS, MW, SD, SZ, UG, AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG

CA 2253924	AA	19971127	CA 1997-2253924	19970519
------------	----	----------	-----------------	----------

AU 9729071	A1	19971209	AU 1997-29071	19970519
------------	----	----------	---------------	----------

AU 725630	B2	20001019		
-----------	----	----------	--	--

EP 906425	A2	19990407	EP 1997-923204	19970519
-----------	----	----------	----------------	----------

R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI, RO

BR 9709101	A	19990803	BR 1997-9101	19970519
------------	---	----------	--------------	----------

CN 1225683	A	19990811	CN 1997-196317	19970519
------------	---	----------	----------------	----------

NZ 332648	A	20000526	NZ 1997-332648	19970519
-----------	---	----------	----------------	----------

JP 2000512489	T2	20000926	JP 1997-541799	19970519
---------------	----	----------	----------------	----------

PRIORITY APPLN. INFO.:			GB 1996-10484	A	19960518
------------------------	--	--	---------------	---	----------

			GB 1997-7844	A	19970418
--	--	--	--------------	---	----------

			WO 1997-GB1372	W	19970519
--	--	--	----------------	---	----------

AB Vasoactive amine binding proteins (VABPs) are provided that specifically bind to vasoactive amines with a dissociation constant of  $<10^{-7}$  M and which belong to the same protein family as MS-HBP1,

**THIS PAGE BLANK (USPTO)**



FS-HBP1, FS-HBP2 and D.

RET6. Thus, 4 VASPs were isolated, and their cDNAs cloned and sequenced, from ticks: FS-HBP1 (female-specific histamine-binding protein 1), FS-HBP2 (female-specific histamine-binding protein 2), and MS-HBP1 (male-specific histamine-binding protein 1) from *Rhipicephalus appendiculatus*; and D.RET6 from *Dermacenter reticularis*. These 4 VASPs possess similar amino acid sequences and predicted secondary structures. The VASPs bind histamine in mammals, and can be used as anti-inflammatory agents to regulate histamine action and to control its pathol. effects.

IT 200220-32-2 200220-33-3 200220-34-4  
200220-35-5

RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses) (amino acid sequence; cloning and functions of vasoactive amine-binding proteins from ticks)

RN 200220-32-2 HCAPLUS

CN Protein FS-HBP1 (*Rhipicephalus appendiculatus* female-specific histamine-binding) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 200220-33-3 HCAPLUS

CN Protein FS-HBP2 (*Rhipicephalus appendiculatus* female-specific histamine-binding) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 200220-34-4 HCAPLUS

CN Protein MS-HBP1 (*Rhipicephalus appendiculatus* male-specific histamine-binding) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

RN 200220-35-5 HCAPLUS

CN Protein D.RET6 (*Dermacenter reticularis* histamine-binding) (9CI) (CA INDEX NAME)

\*\*\* STRUCTURE DIAGRAM IS NOT AVAILABLE \*\*\*

IC ICM C12N015-12

ICS C12N015-86; A01K067-027; A61K038-17; C12N005-10; G01N033-68; C12N015-62; C07K014-435

CC 2-8 (Mammalian Hormones)

Section cross-reference(s): 1, 3

ST vasoactive amine binding protein tick; histamine binding protein tick; *Rhipicephalus* vasoactive amine binding protein; *Dermacenter* vasoactive amine binding protein; sequence vasoactive binding protein cDNA tick; inflammation inhibitor VASP protein tick

IT Proteins, specific or class

RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses) (D.RET6; cloning and functions of vasoactive amine-binding proteins from ticks)

IT Proteins, specific or class

RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses) (FS-HBP1 (female-specific histamine-binding protein 1); cloning and functions

**THIS PAGE BLANK (USPTO)**

- of vasoactive amine-binding proteins from ticks)
- IT Proteins, specific or class
  - RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)
  - (FS-HBP2 (female-specific histamine-binding protein 2); cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Proteins, specific or class
  - RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)
  - (MS-HBP1 (male-specific histamine-binding protein 1); cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Proteins, specific or class
  - RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)
  - (VASP (vasoactive amine-binding protein); cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Amines, biological studies
  - RL: BPR (Biological process); BSU (Biological study, unclassified); BIOL (Biological study); PROC (Process)
  - (biogenic, vasoactive; cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Anti-inflammatory agents
  - Antihistamines
  - Dermacentor reticulatus
  - Mite and Tick
  - Molecular cloning
  - Rhipicephalus appendiculatus
  - Scorpion
  - Snake
  - Spider
  - Transformation, genetic
  - Vaccines
  - (cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Fusion proteins (chimeric proteins)
  - RL: BPN (Biosynthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)
  - (cloning and functions of vasoactive amine-binding proteins from ticks)
- IT Parasite
  - (ecto-; cloning and functions of vasoactive amine-binding proteins from ticks)
- IT cDNA sequences
  - (for vasoactive amine-binding proteins from ticks)
- IT Protein sequences
  - (of vasoactive amine-binding proteins from ticks)
- IT Body fluid
  - (quantification of histamine levels in body fluids; cloning and functions of vasoactive amine-binding proteins from ticks)
- IT 200220-32-2 200220-33-3 200220-34-4 200220-35-5
  - RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)
  - (amino acid sequence; cloning and functions of vasoactive amine-binding proteins from ticks)
- IT 51-45-6, Histamine, biological studies

**THIS PAGE BLANK (USPTO)**

RL: ANT (Analyte); BPR (Biological process); BSU (Biological study, unclassified); ANST (Analytical study); BIOL (Biological study); PROC (Process)

(cloning and functions of vasoactive amine-binding proteins from ticks)

IT 200220-28-6P 200220-29-7P 200220-30-0P 200220-31-1P

RL: BPN (Biosynthetic preparation); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)

(nucleotide sequence; cloning and functions of vasoactive amine-binding proteins from ticks)

**THIS PAGE BLANK (USPTO)**

(nucleotide sequence; high affinity serotonin- and histamine-binding lipocalin sequence and identification from tick saliva)

REFERENCE COUNT: 25 THERE ARE 25 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L32 ANSWER 4 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2001:8145 HCAPLUS

DOCUMENT NUMBER: 134:128741

TITLE: Nitrophorins and related antihemostatic lipocalins from *Rhodnius prolixus* and other blood-sucking arthropods

AUTHOR(S): Montfort, William R.; Weichsel, Andrzej; Andersen, John F.

CORPORATE SOURCE: Department of Biochemistry, University of Arizona, Tucson, AZ, 85721, USA

SOURCE: Biochimica et Biophysica Acta (2000), 1482(1-2), 110-118

CODEN: BBACAQ; ISSN: 0006-3002

PUBLISHER: Elsevier Science B.V.

DOCUMENT TYPE: Journal; General Review

LANGUAGE: English

AB A review, with 53 refs. Recent gene sequence and crystal structure detns. of salivary proteins from several blood-sucking arthropods have revealed an unusual evolutionary relationship: many such proteins derive their functions from lipocalin protein folds. Many blood-sucking arthropods have independently evolved the ability to overcome a host organism's means of preventing blood loss (called hemostasis). Most blood feeders have proteins that induce vasodilation, inhibit blood coagulation, and reduce inflammation, but do so by distinctly different mechanisms. Despite this diversity, in many cases the antihemostatic activities in such organisms reside in proteins with lipocalin folds. Thirteen such lipocalins are described in this review, with a particular focus on the heme-contg. nitrophorins from *R. prolixus*, which transport nitric oxide, sequester histamine, and disrupt blood coagulation. Also described are the anti-platelet compds. RPAI, moubatin, and pallidipin from *R. prolixus*, *Ornithodoros moubata*, and *Triatoma pallidipennis*; the antithrombin protein triabin from *T. pallidipennis*; and the tick histamine binding proteins from *Rhipicephalus appendiculatus*.

CC 12-0 (Nonmammalian Biochemistry)

Section cross-reference(s): 6

ST review blood sucking arthropod nitrophorin lipocalin

IT Arthropod (Arthropoda)

(blood-sucking; nitrophorins and related antihemostatic lipocalins from *Rhodnius prolixus* and other blood-sucking arthropods)

IT Proteins, specific or class

RL: BAC (Biological activity or effector, except adverse); BOC (Biological occurrence); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); OCCU (Occurrence)

(lipocalins; nitrophorins and related antihemostatic lipocalins from *Rhodnius prolixus* and other blood-sucking arthropods)

IT *Rhodnius prolixus*

Salivary gland

(nitrophorins and related antihemostatic lipocalins from *Rhodnius prolixus* and other blood-sucking arthropods)

IT Proteins, specific or class

RL: BAC (Biological activity or effector, except adverse); BOC (Biological occurrence); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); OCCU (Occurrence)

(nitrophorins; nitrophorins and related antihemostatic lipocalins from *Rhodnius prolixus* and other blood-sucking arthropods)

**THIS PAGE BLANK (USPTO)**



REFERENCE COUNT: 53 THERE ARE 53 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L32 ANSWER 5 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2001:8143 HCAPLUS

DOCUMENT NUMBER: 134:128740

TITLE: Tick **histamine-binding proteins**: lipocalins with a second binding cavity

AUTHOR(S): Paesen, Guido C.; Adams, Peter L.; Nuttall, Patricia A.; Stuart, David L.

CORPORATE SOURCE: CEH Institute of Virology and Environmental Microbiology, Oxford, OX1 3SR, UK

SOURCE: Biochimica et Biophysica Acta (2000), 1482(1-2), 92-101

CODEN: BBACAQ; ISSN: 0006-3002

PUBLISHER: Elsevier Science B.V.

DOCUMENT TYPE: Journal; General Review

LANGUAGE: English

AB A review, with 26 refs. Tick **histamine-binding proteins** (HBPs) are lipocalins with 2 binding pockets. One of these binds histamine with a high affinity and is found at the position expected from other lipocalins, adjacent to the .OMEGA.-loop at the open-end of the .beta.-barrel. A 2nd binding cavity, which is a low-affinity site for histamine in 1 of the HBPs, is located at the end of the barrel that is closed off in other lipocalins. To create the 2nd site, the closed-end region has undergone a major reconstruction. Typical lipocalin characteristics, such as the 310 helix and a structural cluster of highly conserved residues, have been lost, while an .alpha.-helix now shields the cavity from the exterior. The prominence of acidic residues in the binding pockets is another distinctive characteristic of HBPs. Whereas most lipocalins have highly hydrophobic binding cavities designed to bind lipophilic compds., HBPs have evolved to trap cationic, hydrophilic mols.

CC 12-0 (Nonmammalian Biochemistry)

Section cross-reference(s): 6

ST review tick **histamine binding protein**

IT Proteins, specific or class

RL: BOC (Biological occurrence); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); OCCU (Occurrence)  
(histamine-binding; tick **histamine-binding proteins**)

IT Mite and Tick

(tick **histamine-binding proteins**)

REFERENCE COUNT: 26 THERE ARE 26 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L32 ANSWER 6 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1994:236350 HCAPLUS

DOCUMENT NUMBER: 120:236350

TITLE: Purification of histamine receptor proteins from detergent-solubilized human peripheral blood mononuclear cells

AUTHOR(S): Warlow, Robert S.; Rajasekariah, Poornima; Lambert, Paul; Morgan, Janelle; Dao, Lan Phuong; Bernard, Claude C. A.; Walls, Ronald S.

CORPORATE SOURCE: Immunology Department, Concord Repatriation Hospital, Sydney, Australia

SOURCE: Biochemistry (1994), 33(16), 4800-11  
CODEN: BICHAW; ISSN: 0006-2960

**THIS PAGE BLANK (USPTO)**

DOCUMENT TYPE: Journal  
 LANGUAGE: English

- AB Histamine is released from mast cells and basophils by either immunol. or nonimmunol. mechanisms. Histamine, which is the most potent short acting mediator released from these cells, exerts its diverse biol. actions by binding to cell surface histamine receptors. The authors report the affinity purifn. of histamine receptor proteins from Triton X-100 solubilized peripheral human blood mononuclear cells which include lymphocytes and monocytes. Three different designs of histamine affinity columns were constructed; all three resulted in the same material being eluted. This consisted of bands which on SDS-PAGE after boiling and redn. had the following mol. wts.: 193K, 84K, 58K, 48K, 37K, and 16K. The most abundant bands were of mol. wts. 193K, 48K, and 16K, and these were disulfide bonded together to form a high mol. wt. complex. The 58K band was present in lower amts. than the others, and in only a few fractions. It had the same mol. wt. as the dimeric form of histamine methyltransferase which is present in small amts. in mononuclear cells and may therefore have copurified. The **histamine binding proteins** described in this report were purified by conventional affinity chromatog., rather than by an expression cloning approach which obviates the use of any protein chem. Consequently, the authors had the advantage of being able to verify the histamine binding specificity of the purified proteins directly and with several independent assays as follows. The histamine binding specificity of all three columns was established by specific elution with histamine, by preabsorption of crude cell ext. with excess free histamine prior to column application, and by comparison with control columns. Independent detn. of the binding specificity, using a radioreceptor dot blot assay, of the eluate contg. only the 193K, 48K, and 16K disulfide-linked subunits confirmed that the purified material bound specifically to [3H]histamine and that a 300-500-fold degree of purifn. from tissue ext. had been obtained. Following cell surface radioreceptor crosslinking of radiolabeled histamine to intact mononuclear cells, the 16K band was detected, indicating it to be the ligand-binding subunit for histamine. These same three proteins were purified from T lymphocyte and monocytoïd cell lines, indicating that both lymphocyte and monocyte subsets of mononuclear cells express these proteins. The trimol. structure, consisting of 193K, 48K, and 16K subunits appears to be a novel histamine cell surface receptor protein complex as the mol. wts. of the 193K, 48K, and 16K bands bore no relation to the predicted mol. wts. of the recently expression cloned H1 receptor gene derived from bovine adrenal and H2 receptor gene derived from human parietal cells; this may reflect receptor heterogeneity within a tissue, between tissues, and/or between species.
- CC 2-2 (Mammalian Hormones)
- ST histamine receptor mononuclear cell purifn characterization; affinity chromatog histamine receptor mononuclear cell; lymphocyte histamine receptor purifn characterization; monocyte histamine receptor purifn characterization; structure histamine receptor mononuclear cell
- IT Monocyte  
 (histamine receptors of human, purifn. and characterization of)
- IT Cell membrane  
 (histamine receptors of, of human blood mononuclear cells, purifn. and characterization of)
- IT Conformation and Conformers  
 Disulfide group  
 (of histamine receptors, of human blood mononuclear cells)
- IT Lymphocyte  
 (T-cell, histamine receptors of human, purifn. and characterization of)
- IT Chromatography, column and liquid  
 (affinity, of histamine receptors, of human blood mononuclear cells)

**THIS PAGE BLANK (USPTO)**

- IT Receptors  
RL: BIOL (Biological study)  
(histaminic, of blood mononuclear cells, of human, purifn. and characterization of)
- IT Leukocyte  
(mononuclear, histamine receptors of human, purifn. and characterization of)
- IT 51-45-6P, Histamine, biological studies  
RL: BIOL (Biological study); PREP (Preparation)  
(receptor binding of, after purifn. from human blood mononuclear cells)

L32 ANSWER 7 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1991:421847 HCAPLUS  
DOCUMENT NUMBER: 115:21847  
TITLE: Histaglobulin and changes in conjunctival structure in vernal conjunctivitis  
AUTHOR(S): Volokhovskaya, Z. P.; Kadyrova, M. Ya.  
CORPORATE SOURCE: Turk. NII Glazn. Bolezn., Ashkhabad, USSR  
SOURCE: Zdravookhranenie Turkmenistana (1990), (10), 15-18  
CODEN: ZDTUAB; ISSN: 0513-8736  
DOCUMENT TYPE: Journal  
LANGUAGE: Russian

- AB The histochem. effects of 2-3-wk course treatment with histaglobin (histamine-binding globulin) on eye conjunctiva were studied in humans suffering from spring allergic conjunctivitis. The treatment decreased infiltration with basophils responsible for amine release and allergic symptoms.
- CC 1-7 (Pharmacology)  
Section cross-reference(s): 15
- ST allergic conjunctivitis histochem histaglobulin histamine binding;  
**histamine binding protein** allergic conjunctivitis histochem
- IT Proteins, specific or class  
RL: BIOL (Biological study)  
(histamine-binding, allergic spring conjunctivitis histochem. response to treatment with, in humans)
- IT Eye, disease or disorder  
(vernal conjunctivitis, histaglobin treatment effects on histochem. of, in humans)

L32 ANSWER 8 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1981:404871 HCAPLUS  
DOCUMENT NUMBER: 95:4871  
TITLE: **Histamine binding proteins**  
separated from human serums by the chromatographic method  
AUTHOR(S): Chachaj, Wladyslaw; Bartecka, Zuzanna; Malolepszy, Jozef  
CORPORATE SOURCE: Dep. Intern. Dis., Inst. Intern. Dis., Wroclaw, Pol.  
SOURCE: Archivum Immunologiae et Therapiae Experimentalis (1980), 28(6), 947-51  
CODEN: AITEAT; ISSN: 0004-069X  
DOCUMENT TYPE: Journal  
LANGUAGE: English

- AB Human serum proteins were bound on and eluted from Sepharose-polylysine-histamine column. Thus obtained protein fraction was referred to as HBP (**histamine-binding protein**). HBPs were examd. for the ability to bind histamine by biol. method applied to isolated guinea pig intestine. Chromatog. sepns. were run on DEAE cellulose column to obtain HBPs which are directly responsible for histamine-binding

**THIS PAGE BLANK (USPTO)**

activity. The results suggest that human serum contains 3 HBP fractions. One of them was identified as orosomucoid, and the 2 remaining ones seem to be glycoproteins belonging to .alpha.1 globulin group.

CC 15-1 (Immunochemistry)  
ST histamine binding protein chromatog  
IT Blood serum  
(histamine-binding proteins of)  
IT Allergy  
(histamine-binding proteins of blood  
serum in relation to)  
IT 51-45-6, biological studies  
RL: BIOL (Biological study)  
(proteins binding, of blood serum)

L32 ANSWER 9 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1979:536384 HCAPLUS  
DOCUMENT NUMBER: 91:136384  
TITLE: Application of affinity chromatography in binding of  
histamine to serum proteins  
AUTHOR(S): Bartecka, Z.; Chachaj, W.; Malepszy, J.  
CORPORATE SOURCE: Inst. Intern. Med., Med. Acad. Wroclaw, Wroclaw, Pol.  
SOURCE: Agents and Actions (1979), 9(1), 35-7  
CODEN: AGACBH; ISSN: 0065-4299  
DOCUMENT TYPE: Journal  
LANGUAGE: English

AB Serum histamine (I)-binding proteins (HBP) were isolated by affinity chromatog. on polylysine-Sepharose 4B-immobilized histamine (II) and characterized by electrophoresis. Dialyzed serum proteins were incubated with a suspension of carrier II for 2 h at 37.degree., and the carrier then was packed in a column, washed with 0.9% NaCl to remove unbound proteins, and eluted with 0.2M EDTA and 6M urea to remove HBP. Polyacrylamide gel electrophoresis in the presence of 0.2% Na dodecyl sulfate sepd. 8 HBP fractions with mol. wts. in the range 80,000-350,000. A bioassay with isolated guinea pig intestine was used to det. I and to study its inactivation at various HBP concns.

CC 9-2 (Biochemical Methods)  
ST serum histamine binding protein sepn;  
affinity chromatog histamine binding protein  
IT Blood analysis  
(histamine-binding proteins affinity.  
chromatog. in)  
IT Chromatography, column and liquid  
(affinity, of histamine-binding proteins)  
IT Proteins  
RL: PROC (Process)  
(histamine-binding, affinity chromatog. of, of blood serum)  
IT 9012-36-6D, reaction product with poly-L-lysine  
RL: ANST (Analytical study)  
(histamine immobilization on, for histamine-binding  
proteins affinity chromatog.)  
IT 51-45-6, biological studies  
RL: BIOL (Biological study)  
(proteins binding, affinity chromatog. of, of blood serum)

L32 ANSWER 10 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1978:35352 HCAPLUS  
DOCUMENT NUMBER: 88:35352  
TITLE: Histamine and serotonin content in patients with newly  
identified tuberculosis of the lungs complicated by  
nonspecific endobronchitis

**THIS PAGE BLANK (USPTO)**



AUTHOR(S): Shesterina, M. V.; Mikhailova, Yu. V.; Andrzheyuk, N. I.; Elistratova, N. A.  
 CORPORATE SOURCE: Mosk. Nauchno-Issled. Inst. Tuberk., Moscow, USSR  
 SOURCE: Problemy Tuberkuleza (1977), (9), 36-9  
 CODEN: PRTUAX; ISSN: 0032-9533  
 DOCUMENT TYPE: Journal  
 LANGUAGE: Russian  
 AB In 50 adult patients with tuberculosis and nonspecific endobronchitis, the blood levels of histamine and serotonin increased and those of **histamine-binding protein** decreased 2-, 1.5- and 5-fold, resp., in the patients without microflora in the bronchi and 1.4-, 1.5-, and 2-fold, resp., in the patients with addnl. microflora in the bronchi as compared with those of normal subjects. After treatment with antibiotics and antihistamines or corticosteroids, the levels usually normalized.  
 CC 14-1 (Mammalian Pathological Biochemistry)  
 ST histamine serotonin blood tuberculosis; bronchitis microorganism  
 IT tuberculosis; protein serum histamine binding endobronchitis  
 IT Tuberculosis  
 (blood compn. in, endobronchitis in relation to)  
 IT Proteins  
 RL: BIOL (Biological study)  
 (histamine-binding, of blood in tuberculosis and endobronchitis)  
 IT Bronchi  
 (disease, endobronchitis, blood compn. in tuberculosis in relation to)  
 IT 50-67-9, biological studies 51-45-6, biological studies  
 RL: BIOL (Biological study)  
 (of blood, in tuberculosis and endobronchitis)

L32 ANSWER 11 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN  
 ACCESSION NUMBER: 1975:81859 HCAPLUS  
 DOCUMENT NUMBER: 82:81859  
 TITLE: Histamine-serum protein bond  
 AUTHOR(S): Kochman, S.; Jalouzet, R.; Maujean, A.; Cazabat, A.; Dubois de Montreynaud, J. M.  
 CORPORATE SOURCE: Serv. Pneumophysiologie, Cent. Hosp. Univ., Reims, Fr.  
 SOURCE: Nouvelle Presse Medicale (1974), 3(30), 1883  
 CODEN: NPMDAD; ISSN: 0301-1518  
 DOCUMENT TYPE: Journal  
 LANGUAGE: French  
 AB Gel chromatog. of normal human serum on a histamine-fixed, CNBr-activated Sepharose 4B column adsorbed the histaminoplexig compd. of the serum. Elution with 6M guanidine-HCl yielded a protein (max. uv absorbance at 280 nm). Na dodecyl sulfate-polyacrylamide gel electrophoresis showed only 2 bands. Thus, the presence of a **histamine-binding protein** in blood serum is confirmed.  
 CC 6-3 (General Biochemistry)  
 ST blood serum protein histamine binding  
 IT Proteins  
 RL: BIOL (Biological study)  
 (blood-serum, histamine binding by)  
 IT 51-45-6, biological studies  
 RL: BIOL (Biological study)  
 (protein binding, of blood serum)

L32 ANSWER 12 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN  
 ACCESSION NUMBER: 1974:502483 HCAPLUS  
 DOCUMENT NUMBER: 81:102483  
 TITLE: New aspects of histaminopexy  
 AUTHOR(S): Lonovics, J.; Gecse, A.; Karady, S.

**THIS PAGE BLANK (USPTO)**

CORPORATE SOURCE: Inst. Pathophysiol., Univ. Sch. Med., Szeged, Hung.  
 SOURCE: Archives Internationales de Physiologie et de  
 Biochimie (1974), 82(1), 41-7  
 CODEN: AIPBAY; ISSN: 0003-9799

DOCUMENT TYPE: Journal  
 LANGUAGE: English

AB The mechanism of formation and release of histaminopexic peptide was studied in the rat. Detns. were carried out with plasma and serum which were first subjected to gel filtration on Sephadex G-25. Agarose gel electrophoresis was used to sep. polypeptides, and free histamine (I) was assayed with an atropinized guinea pig ileum suspended in Tyrode's soln. Whereas I-binding activity of plasma was in the protein fraction, in serum it was in the peptide region; after papain digestion, results with plasma were similar to those with serum. Adrenalectomy caused loss of I-binding activity, unless rats were treated with prednisolone succinate. The I-binding peptide probably released from albumin by a papainlike action during blood coagulation. Adrenal cortex function is needed for maintenance of I-binding activity.

CC 13-13 (Mammalian Biochemistry)

ST histamine binding protein blood; adrenal  
 histamine binding protein

IT Proteins

RL: BIOL (Biological study)  
 (histamine binding, blood-plasma, regulation of)

IT Corticosteroids, biological studies

RL: BIOL (Biological study)  
 (histamine-binding peptide of blood serum in response to)

IT Peptides, biological studies

RL: BIOL (Biological study)  
 (histamine-binding, of blood serum, regulation of)

IT 2920-86-7

RL: BIOL (Biological study)  
 (histamine-binding peptide of blood serum in response to)

IT 51-45-6, biological studies

RL: BIOL (Biological study)  
 (peptide of blood serum binding of, regulation of)

L32 ANSWER 13 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1972:111092 HCAPLUS

DOCUMENT NUMBER: 76:111092

TITLE: Histamine-binding property of human serum

AUTHOR(S): Gecse, A.; Lonovics, J.; Szekeres, L.; Zsilinszky, E.;  
 West, G. B.

CORPORATE SOURCE: Sch. Med., Univ. Szeged, Szeged, Hung.

SOURCE: Journal of Pharmacy and Pharmacology (1972), 24(1),  
 70-1

CODEN: JPPMAB; ISSN: 0022-3573

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Expts. performed on serum from healthy adult humans indicated that the albumin fraction in some way binds histamine, whereas all other fractions of the plasma were lacking in such activity. When the fraction contg. 96% of albumin was digested with papain, and then passed through Sephadex G-25, the histamine-binding activity moved to the polypeptide region of the plasma. All the low-mol.-wt. polypeptide portion had been released from the albumin during coagulation, or by papain digestion. Plasma albumin or serum peptide from blood of 8 allergic patients (generalized dermatoses) lacked histamine-binding activity, indicating that lack of a histaminopexic substance may be involved in the pathogenesis of allergic diseases.

**THIS PAGE BLANK (USPTO)**

CC 13 (Mammalian Biochemistry)  
 ST blood histamine binding protein  
 IT Allergy  
     (histamine binding by albumins of blood serum in)  
 IT Albumins, blood serum  
     RL: BIOL (Biological study)  
     (histamine binding by, in allergy)  
 IT 51-45-6, biological studies  
     RL: BIOL (Biological study)  
     (albumins of blood serum binding of, in allergy)

L32 ANSWER 14 OF 14 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1968:75101 HCAPLUS

DOCUMENT NUMBER: 68:75101

TITLE: Specificity effects of cation binding to the macroion  
     in the bovine serum albumin-histamine interaction

AUTHOR(S): Botre, Claudio; Marchetti, Marcello; Borghi, Saverio

CORPORATE SOURCE: Univ. Rome, Rome, Italy

SOURCE: Biochimica et Biophysica Acta (1968), 154(2), 360-6

CODEN: BBACAQ; ISSN: 0006-3002

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The specific interaction between bovine serum albumin and histamine and the influence of the ionic medium on this interaction were studied. This interaction is dependent on ionic strength. Evidence is given for a crit. saline concn. which can be regarded as a threshold defining 2 sharply different behaviors of bovine serum albumin in the presence of histamine. Potentiometric measurements carried out by means of coupled perm-selective membrane electrodes provided data which showed that for a given ionic strength, bovine serum albumin requires an amt. of histamine which varies with the nature of the cations. Alkali binding to the bovine serum albumin macroion in the presence of histamine does not follow the sequence expected for polycarboxylic acids, since K<sup>+</sup> proved to be the most tightly bound. 20 references.

CC 2 (General Biochemistry)

ST BINDING; HISTAMINE BINDING PROTEINS;  
 ELECTROLYTES ALBUMINS HISTAMINE; CATIONS ALBUMINS HISTAMINE BINDING;  
 ALBUMINS HISTAMINE BINDING; PROTEINS HISTAMINE BINDING

IT Cations, biological studies  
     (in histamine reaction with albumins)

IT Albumins, blood serum  
     RL: RCT (Reactant); RACT (Reactant or reagent)  
     (reaction of, with histamine, cation binding effect on)

IT 51-45-6, reactions  
     RL: RCT (Reactant); RACT (Reactant or reagent)  
     (with albumin, cation binding effect on)

**THIS PAGE BLANK (USPTO)**



XX Claim 4; Page 5-6; 19pp; English.  
 PS The present invention relates to the use of a histacalin protein  
 CC for treating or preventing conjunctivitis. The present sequence  
 CC is the histacalin protein FS-HBp1. The invention is particularly useful  
 CC in the treatment of allergic or seasonal conjunctivitis.  
 XX  
 SO Sequence 172 AA;  
 Query Match 100.0%; Score 952; DB 22; Length 172;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-94;  
 Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DKPWADAEANGEHQDAMKHLQKLYEENYDLIKATYKNDPVWGNDFTCVGTAAQNLNDE 60  
 DB 1 DKPWADAEANGEHQDAMKHLQKLYEENYDLIKATYKNDPVWGNDFTCVGTAAQNLNDE 60  
 QY 61 KNVEAMFEMNNADTVYQHFFEKATPDKMTGYKNENAIYQTEDEGQVLTDLAFSDNDCY 120  
 DB 61 KNVEAMFEMNNADTVYQHFFEKATPDKMTGYKNENAIYQTEDEGQVLTDLAFSDNDCY 120  
 QY 121 VYVAGPDGSGAGYELMATDYTDVPASCLEKFNKYAGLPVRYVTSDCPE 172  
 DB 121 VYVAGPDGSGAGYELMATDYTDVPASCLEKFNKYAGLPVRYVTSDCPE 172  
 RESULT 2  
 AAB73260  
 ID AAB73260 standard; protein; 172 AA.  
 AC AAB73260;  
 DT 15-MAY-2001 (first entry)  
 DE Histacalin protein FS-HBp1.  
 KW FS-HBp1; histacalin protein; antinflammatory; antiallergic;  
 KM ophthalmological; allergic rhinitis.  
 XX  
 OS Unidentified.  
 PN WO200116164-A2.  
 PD 08-MAR-2001.  
 PF 24-AUG-2000; 2000WO-GB03287.  
 PR 01-SEP-1999; 99GB-0020673.  
 PA (EVOLOC) EVOLUTEC LTD.  
 PI Nuttall PA, Paesen GC;  
 DR WPI; 2001-218521/22.  
 XX  
 PT Use of histacalin proteins for treating or preventing allergic  
 PT rhinitis, or for manufacturing a medicament for treating or preventing  
 PT allergic rhinitis, e.g. seasonal or perennial allergic rhinitis  
 PS Disclosure; Pages 4-6; 19pp; English.  
 XX  
 CC The present invention relates to a method for treating or preventing  
 CC allergic rhinitis. The method involves employing a blood-feeding  
 CC ectoparasite-derived (e.g. tick-derived) histacalin protein. The present  
 CC sequence is one such histacalin protein. The histacalin protein, is  
 CC useful for treating or preventing allergic rhinitis, both seasonal and  
 CC perennial allergic conjunctivitis.  
 XX  
 SO Sequence 172 AA;  
 Query Match 100.0%; Score 952; DB 22; Length 172;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-94;

Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DKPWADAEANGEHQDAMKHLQKLYEENYDLIKATYKNDPVWGNDFTCVGTAAQNLNDE 60  
 DB 1 DKPWADAEANGEHQDAMKHLQKLYEENYDLIKATYKNDPVWGNDFTCVGTAAQNLNDE 60  
 QY 61 KNVEAMFEMNNADTVYQHFFEKATPDKMTGYKNENAIYQTEDEGQVLTDLAFSDNDCY 120  
 DB 61 KNVEAMFEMNNADTVYQHFFEKATPDKMTGYKNENAIYQTEDEGQVLTDLAFSDNDCY 120  
 QY 121 VYVAGPDGSGAGYELMATDYTDVPASCLEKFNKYAGLPVRYVTSDCPE 172  
 DB 121 VYVAGPDGSGAGYELMATDYTDVPASCLEKFNKYAGLPVRYVTSDCPE 172  
 RESULT 3  
 AAM37446  
 ID AAM37446 standard; Protein; 190 AA.  
 AC AAM37446;  
 DT 08-JUN-1998 (first entry)  
 DE Tick vasoactive amine binding protein 1 FS-HBp1.  
 KW Female-specific vasoactive amine binding protein 1; FS-HCp1;  
 KW histamine; serotonin; assay; antihistamine; anti-inflammatory;  
 KW insect bite; snake bite; scorpion bite; dermatitis; vaccine;  
 KM transgenic animal; tick.  
 XX  
 OS Rhipicephalus appendiculatus.  
 PN Key location/Qualifiers  
 PF Peptide 1..18  
 FT /label= sig\_peptide  
 XX  
 PN WO9744451-A2.  
 PD 27-NOV-1997.  
 PF 19-MAY-1997; 97MO-GB01372.  
 PR 18-APR-1997; 97GB-0007844.  
 PR 18-MAY-1996; 96GB-0010484.  
 PA (OXFO-) OXFORD VACS LTD.  
 PI Nuttall PA, Paesen GC;  
 DR WPI; 1998-018506/02.  
 DR N-PSDB; AAV00227.  
 XX  
 PT New vasoactive amine binding proteins and related nucleic acid,  
 PT vectors - transformed cells and transgenic animals, used for  
 PT assaying or removing histamine and as antihistamine or  
 PT anti-inflammatory agents  
 PS Example 2; Fig 1; 44pp; English.  
 XX  
 CC This protein comprises tick Rhipicephalus appendiculatus (Ra) novel  
 CC female-specific histamine binding protein 1 (FS-HBp1). Its amino  
 CC acid sequence was deduced from a cDNA clone (see AAV00227) obtained  
 CC from a salivary gland cDNA library. FS-HSP2 and male-specific  
 CC HBp1 (see AAM37447-48) and a related protein, D.RET6 (see AAM37449)  
 CC from Dermacentor reticulatus, were also identified. These novel  
 CC vasoactive amine binding proteins (VABPs) can be expressed in  
 CC host cells using e.g. a baculovirus expression system. They can  
 CC be used: (i) to assay histamine (or other VA such as serotonin) in  
 CC body fluids or cell culture supernatants, e.g. to monitor the  
 CC effect of allergens; (ii) for binding VA, e.g. to remove histamine  
 CC from blood, food, cell cultures etc.; (iii) as an antihistamine or  
 CC anti-inflammatory agents, e.g. for treating insect, snake or  
 CC scorpion bites or dermatitis, or as a carrier for slow release of



CC histamine-related compounds; (iv) in vaccines to protect against  
CC metazoan parasites, especially in animals; (v) as reagents for the  
CC studying inflammation, involvement of VA in ulcer formation or the  
CC immune response etc. VABPs provide a more sensitive assay for  
CC histamine than low-affinity antibodies currently used. They may  
CC also be more effective and safer than conventional antihistamines.

XX Sequence 190 AA;

Query Match 100.0%; Score 952; DB 19; Length 190;  
Best Local Similarity 100.0%; Pred. No. 4.5e-94;  
Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DKPVMDEANGEHQDAMKHLQKLYEENYDLIKATYRNDPVWGNDFCVGTAQNLNDE 60

DB 19 DKPVMDEANGEHQDAMKHLQKLYEENYDLIKATYRNDPVWGNDFCVGTAQNLNDE 78

QY 61 KNEVAFMFNNMADTVYQHTEFEKATPDKMYGYNKENAITTYQTEGQVLTDLAFSDNDCY 120

DB 79 KNEVAFMFNNMADTVYQHTEFEKATPDKMYGYNKENAITTYQTEGQVLTDLAFSDNDCY 138

QY 121 VIVALGPDGSGAGYELMATDVTDPVASCLEKFNENYAAGLPVRDVTSDCLPE 172

DB 139 VIVALGPDGSGAGYELMATDVTDPVASCLEKFNENYAAGLPVRDVTSDCLPE 190

#### RESULT 4

AAV18078 standard; protein: 190 AA.

AAV18078;

DT 06-AUG-1999 (first entry)

DE Histamine binding protein FS-HBP1.

XX Histamine binding protein; serotonin binding compound; inflammation;

KW gastric acid secretion; allergy; type I hypersensitivity reaction;

KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;

KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;

KW respiratory disease; coronary heart disease; cellular growth regulator;

KW tissue repair; blood-sucking ectoparasite; therapy.

XX Rhinopneumonius appendiculatus.

XX W09927104-A1.

PD 03-JUN-1999.

XX 26-NOV-1998; 98MO-G803530.

XX 26-JUN-1998; 98GB-0013917.

XX 26-NOV-1997; 97GB-0025046.

PA (OXFO-) OXFORD VACS LTD.

PI Nuttall PA, Paesen GC;

XX WPI: 1999-357841/30.

DR N-PSDB; AAX76964.

XX Histamine and serotonin binding compounds useful for the treatment

PT of allergies

XX Claim 1; Fig 1; 84pp; English.

CC This sequence is an example of a histamine or serotonin binding  
CC compound (A), of the invention. The compounds are useful for regulating  
CC the action of histamine and serotonin (in e.g. inflammation and gastric  
CC acid secretion), the detection, quantification and removal of histamine  
CC or serotonin (in animals, plants, cell cultures, food materials, or  
CC humans) and in the treatment of various diseases and allergies  
CC (e.g. type I hypersensitivity reactions, urticaria, asthma, allergic

CC rhinitis (hay fever), atopic dermatitis, insect bites and food and drug  
CC allergies, abnormal blood pressure, migraine, psychological disorders,  
CC respiratory disease, and coronary heart disease). Histamine may also be  
CC used to regulate cellular growth and tissue repair. The molecules may  
CC also be used as components of vaccines directed against blood-sucking  
CC ectoparasites.

XX Sequence 190 AA;

Query Match 100.0%; Score 952; DB 20; Length 190;  
Best Local Similarity 100.0%; Pred. No. 4.5e-94;  
Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DKPVMDEANGEHQDAMKHLQKLYEENYDLIKATYRNDPVWGNDFCVGTAQNLNDE 60

DB 19 DKPVMDEANGEHQDAMKHLQKLYEENYDLIKATYRNDPVWGNDFCVGTAQNLNDE 78

QY 61 KNEVAFMFNNMADTVYQHTEFEKATPDKMYGYNKENAITTYQTEGQVLTDLAFSDNDCY 120

DB 79 KNEVAFMFNNMADTVYQHTEFEKATPDKMYGYNKENAITTYQTEGQVLTDLAFSDNDCY 138

QY 121 VIVALGPDGSGAGYELMATDVTDPVASCLEKFNENYAAGLPVRDVTSDCLPE 172

DB 139 VIVALGPDGSGAGYELMATDVTDPVASCLEKFNENYAAGLPVRDVTSDCLPE 190

#### RESULT 5

AAB74289 standard; protein: 171 AA.

AAB74289;

DT 20-JUN-2001 (first entry)

DE Histacalin protein FS-HBP2.

XX Histacalin; FS-HBP1, conjunctivitis.

XX Unidentified.

XX W0200115719-A2.

XX 08-MAR-2001.

XX 24-AUG-2000; 2000MO-G803282.

XX 01-SEP-1999; 99GB-0020674.

PA (EVOL-) EVOLUTEC LTD.

PI Nuttall PA, Paesen GC;

XX WPI: 2001-257675/26.

XX Use of histacalin proteins for treating or preventing non-infective

PT conjunctivitis, or for manufacturing a medicament for treating or

PT preventing conjunctivitis, e.g. seasonal or perennial allergic

PT conjunctivitis

XX Claim 4; Page 5-6; 19pp; English.

CC The present invention relates to the use of a histacalin protein

CC for treating or preventing conjunctivitis. The present sequence

CC is the histacalin protein FS-HBP1. The invention is particularly useful

CC in the treatment of allergic or seasonal conjunctivitis.

XX Sequence 171 AA;

Query Match 66.2%; Score 630; DB 22; Length 171;  
Best Local Similarity 66.5%; Pred. No. 1.7e-59;  
Matches 113; Conservative 20; Mismatches 37; Indels 0; Gaps 0;

QY 1 DKPVMDEANGEHQDAMKHLQKLYEENYDLIKATYRNDPVWGNDFCVGTAQNLNDE 60

```

Db      1  NOPDMADEANGAHODAMKSLKADVENVYMKATYKNDPVGNDFTCVGVAMANDVNEDE 60
      61  KNYEAFMFPMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      62  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      63  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      64  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      65  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      66  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      67  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      68  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      69  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      70  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      71  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      72  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      73  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      74  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      75  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      76  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      77  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      78  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      79  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      80  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      81  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      82  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      83  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      84  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      85  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      86  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      87  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      88  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      89  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      90  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      91  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      92  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      93  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      94  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      95  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      96  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      97  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      98  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      99  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      100 KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120

```

RESULT 6  
AAB73261 standard; protein; 171 AA.

AC AAB73261.  
DE 15-MAY-2001 (first entry)  
DE Histacalin protein FS-HBP2.  
KM FS-HBP2; histacalin protein; antinflammatory; antiallergic;  
KM ophthalmological; allergic rhinitis.  
OS Unidentified.  
PN WO200116164-A2.  
PD 08-MAR-2001.  
PF 24-AUG-2000; 2000MO-GB03287.  
PR 01-SEP-1999; 99GB-0020673.  
PA (EVOL-) EVOLUTEC LTD.  
PI Nuttall PA, Paesen GC;  
PT WPI; 2001-218521/22.

Use of histacalin proteins for treating or preventing allergic rhinitis, or for manufacturing a medicament for treating or preventing allergic rhinitis, e.g. seasonal or perennial allergic rhinitis.  
Disclosure; Pages 4-6; 19pp; English.  
The present invention relates to a method for treating or preventing allergic rhinitis. The method involves employing a blood-feeding ectoparasite-derived (e.g. tick-derived) histacalin protein. The present sequence is one such histacalin protein. The histacalin protein, is useful for treating or preventing allergic rhinitis, both seasonal and perennial allergic conjunctivitis.

Sequence 171 AA:

Query Match 66.2%; Score 630; DB 22; Length 171;  
Best Local Similarity 66.5%; Pred. No. 1.7e-59;  
Matches 113; Conservative 20; Mismatches 37; Indels 0; Gaps 0;

```

Db      1  DRVWADDEANGHODAMKSLKADVENVYMKATYKNDPVGNDFTCVGVAMANDVNEDE 60
      61  KNYEAFMFPMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      62  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      63  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      64  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      65  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      66  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      67  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      68  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      69  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      70  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      71  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      72  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      73  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      74  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      75  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      76  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      77  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      78  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      79  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      80  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      81  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      82  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      83  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      84  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      85  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      86  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      87  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      88  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      89  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      90  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      91  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      92  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      93  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      94  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      95  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      96  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      97  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      98  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      99  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      100 KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120

```

RESULT 7  
AAM37447 standard; Protein; 190 AA.  
AC AAM37447.  
DE 08-JUN-1998 (first entry)  
DE Tick vasoactive amine binding protein 2 FS-HBP2.  
KM Female-specific vasoactive amine binding protein 1; FS-HBP1;  
KM histamine; serotonin; assay; antihistamine; anti-inflammatory;  
KM insect bite; snake bite; scorpion bite; dermatitis; vaccine;  
KM transgenic animal; tick.  
OS Rhipicephalus appendiculatus.

Key Location/Qualifiers  
Peptide 1..19  
/label= sig\_peptide  
WO9744451-A2.

27-NOV-1997.  
19-MAY-1997; 97WO-GB01372.  
18-APR-1997; 97GB-0007844.  
18-MAY-1996; 96GB-0010484.

(OXFO-) OXFORD VACS LTD.  
Nuttall PA, Paesen GC;  
WPI; 1998-018506/02.  
N-PSDB; AAV00228.

New vasoactive amine binding proteins and related nucleic acid, vectors - transformed cells and transgenic animals, used for assaying or removing histamine and as antihistamine or anti-inflammatory agents  
Example 2; Fig 2; 4pp; English.

This protein comprises tick Rhipicephalus appendiculatus (Ra) novel female-specific histamine binding protein 2 (FS-HBP2). Its amino acid sequence was deduced from a cDNA clone (see AAV00228) obtained from a salivary gland cDNA library. FS-HBP1 (see AAM37446) and male-specific HSP1 (see AAM37448) and a related protein, D-RET6 (see AAM37449) from Dermacentor reticulatus, were also identified. These novel vasoactive amine binding proteins (VABPs) can be expressed in host cells using e.g. a baculovirus expression system. They can be used: (i) to assay histamine (or other VA such as serotonin) in body fluids or cell culture supernatants, e.g. to monitor the effect of allergens; (ii) for binding VA, e.g. to remove histamine from blood, food, cell cultures etc.; (iii) as an antihistamine or anti-inflammatory agents, e.g. for treating insect, snake or scorpion bites or dermatitis, or as a carrier for slow release of histamine-related compounds; (iv) in vaccines to protect against metaorgan parasites, especially in animals; (v) as reagents for studying inflammation, involvement of VA in ulcer formation or the immune response etc. VABPs provide a more sensitive assay for histamine than low-affinity antibodies currently used. They may also be more effective and safer than conventional antihistamines.

Sequence 190 AA:

Query Match 66.2%; Score 630; DB 19; Length 190;  
Best Local Similarity 66.5%; Pred. No. 2e-59;  
Matches 113; Conservative 20; Mismatches 37; Indels 0; Gaps 0;

```

Db      1  DRVWADDEANGHODAMKSLKADVENVYMKATYKNDPVGNDFTCVGVAMANDVNEDE 60
      61  KNYEAFMFPMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      62  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      63  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      64  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      65  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      66  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      67  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      68  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      69  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      70  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      71  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      72  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      73  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      74  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      75  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      76  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      77  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      78  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      79  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      80  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      81  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      82  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      83  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      84  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      85  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      86  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      87  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      88  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      89  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      90  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      91  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      92  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      93  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      94  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      95  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      96  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      97  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      98  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      99  KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120
      100 KSIQAEFLFMNNADTYOHTFEKATPDKNMGYKRENAITYOTEDGOVLMDVLAFPSDNCY 120

```

DB 20 NOPDWADEAANGAHODANKSLKADVENYVYWKATYKNDPVWGNDFTCVGMANDVNEDE 79  
QY 61 KNVEAMFEMNNADIVYQHTFEKATPDKMKGKKNENATITTCQEDGOVLTDVLAESDDCY 120  
DB 80 KSIQAEFLFEMNNADINMOFATEKTVAVKMGYNRENARREYETEDGOVETDVLAYSDDNCD 139  
QY 121 VYVYALPGDGSAGYELMATDYTDVPASCLEKFNENYAAGLPVRDVTSPCL 170  
DB 140 VYVYGTGNEGEGYELMTTDDYDNIANCINKFNENYAAGRETRDVTSSACL 189

RESULT 8  
AAV18079  
ID AAV18079 standard; protein; 190 AA.

AC AAV18079;

DT 06-AUG-1999 (first entry)

DE Histamine binding protein FS-HBP2.

XX Histamine binding protein; serotonin binding compound; inflammation;  
KM gastric acid secretion; allergy; type I hypersensitivity reaction;  
KM asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
KM drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
KM respiratory disease; coronary heart disease; cellular growth regulator;  
KM tissue repair; blood-sucking ectoparasite; therapy.

XX Rhipicephalus appendiculatus.

XX WO9927104-A1.

XX 03-JUN-1999.

XX 26-NOV-1998; 98WO-GB03530.

XX 26-JUN-1998; 98GB-0013917.

XX 26-NOV-1997; 97GB-0025046.

XX (OXFORD VACS LTD.

XX Nuttall PA, Paesen GC;

XX MPI: 1999-357841/30.

XX N-PSDB; AAX76965.

XX Histamine and serotonin binding compounds useful for the treatment  
PT of allergies

XX Claim 1; Fig 2; 84pp; English.

CC This sequence is an example of a histamine or serotonin binding  
CC compound (A), of the invention. The compounds are useful for regulating  
CC the action of histamine and serotonin (in e.g. inflammation and gastric  
CC acid secretion), the detection, quantification and removal of histamine  
CC or serotonin (in animals, plants, cell cultures, food materials, or  
CC humans) and in the treatment of various diseases and allergies  
CC (e.g. type I hypersensitivity reactions, urticaria, asthma, allergic  
CC rhinitis (hay fever), atopic dermatitis, insect bites and food and drug  
CC allergies, abnormal blood pressure, migraine, psychological disorders,  
CC respiratory disease, and coronary heart disease). Histamine may also be  
CC used to regulate cellular growth and tissue repair. The molecules may  
CC also be used as components of vaccines directed against blood-sucking  
CC ectoparasites.

XX Sequence 190 AA;

QY Query Match 66.2%; Score 630; DB 20; Length 190;

Best Local Similarity 66.5%; Pred. No. 2e-59; 37; Indels 0; Gaps 0;

Matches 113; Conservative 20; Mismatches 37; Indels 0; Gaps 0;  
QY 1 DKPVWADEAANGAHODANKHLOKLVENYDLIKATYKNDPVWGNDFTCVGMANDVNEDE 60  
:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||

DB 20 NOPDWADEAANGAHODANKSLKADVENYVYWKATYKNDPVWGNDFTCVGMANDVNEDE 79  
QY 61 KNVEAMFEMNNADIVYQHTFEKATPDKMKGKKNENATITTCQEDGOVLTDVLAESDDCY 120  
DB 80 KSIQAEFLFEMNNADINMOFATEKTVAVKMGYNRENARREYETEDGOVETDVLAYSDDNCD 139  
QY 121 VYVYALPGDGSAGYELMATDYTDVPASCLEKFNENYAAGLPVRDVTSPCL 170  
DB 140 VYVYGTGNEGEGYELMTTDDYDNIANCINKFNENYAAGRETRDVTSSACL 189

RESULT 9  
AAB74290  
ID AAB74290 standard; protein; 182 AA.

AC AAB74290;

DT 20-JUN-2001 (first entry)

DE Histacalin protein MS-HBP1.

XX Histacalin; MS-HBP1; conjunctivitis.

XX Unidentified.

XX WO200115719-A2.

XX 08-MAR-2001.

XX 24-AUG-2000; 2000WO-GB03282.

XX 01-SEP-1999; 99GB-0020674.

XX (EVOLUTEC LTD.

XX Nuttall PA, Paesen GC;

XX MPI: 2001-257675/26.

XX Use of histacalin proteins for treating or preventing non-infective  
PT conjunctivitis, or for manufacturing a medicament for treating or  
PT preventing conjunctivitis, e.g. seasonal or perennial allergic  
PT conjunctivitis

XX Claim 4; Page 5-6; 19pp; English.

XX The present invention relates to the use of a histacalin protein  
CC for treating or preventing conjunctivitis. The present sequence  
CC is the histacalin protein FS-HBP1. The invention is particularly useful  
CC in the treatment of allergic or seasonal conjunctivitis.

XX Sequence 182 AA;

QY Query Match 28.1%; Score 267.5; DB 22; Length 182;

Best Local Similarity 39.3%; Pred. No. 1.9e-20; 77; Indels 11; Gaps 7;

Matches 70; Conservative 20; Mismatches 77; Indels 11; Gaps 7;

QY 3 PVWADEAANGAHODANKHLOKLVENYDLIKATYKNDPVWGNDFTCVGMANDVNEDE 62

DB 2 PTWANEAKLSGSDODAMKSLDODQNRYYLAQATQTTDGVWGEFFCVSYAEKIG--KKK 59

QY 63 VEAMFEMNNADIVYQHTFEKATPDKMKGKKNENATITTCQED--GOVLTDVLAESD-DNC 119

DB 60 UNATLLYKKNHLLDLSKESHETITVWKATDYTTENGIKETOTGTOTEDVFEVFSQYKNC 119

QY 120 VYVYALPGDGSAG--YELMAT--DYTDVPASCLEKFNENYA--GHPVRDVT--SDCLP 171

DB 120 DVIFPKRGSDGSEGYELMWSEDKIDKIPDCKFTMAFAQOQEKTVNRNVYDSSCKP 177

RESULT 10  
AAB73262  
ID AAB73262 standard; protein; 182 AA.

XX AAB73262;  
 XX 15-MAY-2001 (first entry)  
 DE Histacalin protein MS-HBPI.  
 XX MS-HBPI; histacalin protein; antiinflammatory; antiallergic;  
 KM ophthalmological; allergic rhinitis.  
 XX Unidentified.  
 OS WO200116164-A2.  
 FN 08-MAR-2001.  
 PD 24-AUG-2000; 2000WO-GB03287.  
 XX 01-SEP-1999; 99GB-0020673.  
 XX (EVOLOC) EVOLOC LTD.  
 PA Nuttall PA, Paesen GC;  
 PI WPI; 2001-218521/22.  
 DR Use of histacalin proteins for treating or preventing allergic  
 XX rhinitis, or for manufacturing a medicament for treating or preventing  
 PT allergic rhinitis, e.g. seasonal or perennial allergic rhinitis -  
 PS Disclosure; Pages 4-6; 19pp; English.  
 XX The present invention relates to a method for treating or preventing  
 CC allergic rhinitis. The method involves employing a blood-feeding  
 CC ectoparasite-derived (e.g. tick-derived) histacalin protein. The present  
 CC sequence is one such histacalin protein. The histacalin protein, is  
 CC useful for treating or preventing allergic rhinitis, both seasonal and  
 CC perennial allergic conjunctivitis.  
 XX Sequence 182 AA;  
 SQ  
 Query Match 28.1%; Score 267.5; DB 22; Length 182;  
 Best Local Similarity 39.3%; Pred. No. 1.9e-20;  
 Matches 70; Conservative 20; Mismatches 77; Indels 11; Gaps 7;  
 QY 3 PWADEAANGHODAMKHLQKLEBENDLIKATYKNDPVNGNDFTCVGTAQNLNDEKN 62  
 DB 2 PTWANEAKLGSYODAMKSLQDQDKRYLAQAQTGTVGVEEFTCVSAEKG--KKR 59  
 QY 63 VEAMFEMNADYVYQHTFEKATPDKMYGYNKENATYQTED--GOVLTVDLAFSD-DNC 119  
 DB 60 LNAFTILYKKNHLDKESHETITVWKAAYDTTENGKIKYETQGRTRQTFEDVFVSDYKNC 119  
 QY 120 YVITYALGPDGSGAG-YELMAT-DYTDVPASCLEKFNRYAA---GLPVDYTT-SDCLP 171  
 DB 120 DVLFVPERKSGSDGDEYELWVSEDKIDIPCCKFTMAFYAQOQEKYRNVYTDSSCKP 177  
 RESULT 11  
 AAM37448  
 ID AAM37448 standard; Protein; 200 AA.  
 AC AAM37448;  
 DT 08-JUN-1998 (first entry)  
 XX Tick vasoactive amine binding protein 1 MS-HBPI.  
 XX Male-specific vasoactive amine binding protein 1; MS-HBPI;  
 KW histamine; serotonin; assay; antihistamine; anti-inflammatory;  
 KM insect bite; snake bite; scorpion bite; dermatitis; vaccine;  
 XX transgenic animal; tick.

OS Rhipicephalus appendiculatus.  
 XX Key Location/Qualifiers  
 FH Peptide 1.18  
 FT /label= Sig-peptide  
 FT Modified-site 79..81  
 FT /note= "Asn is N-glycosylated"  
 XX WO9744451-A2.  
 XX 27-NOV-1997.  
 PD 19-MAY-1997; 97WO-GB01372.  
 XX 18-APR-1997; 97GB-0007844.  
 PR 18-MAY-1996; 96GB-0010484.  
 XX (OXFO-) OXFORD VACS LTD.  
 PA Nuttall PA, Paesen GC;  
 PI WPI; 1998-018506/02.  
 DR N-FSDB; AAV00229.  
 XX New vasoactive amine binding proteins and related nucleic acid,  
 PT vectors - transformed cells and transgenic animals, used for  
 PT assaying or removing histamine and as antihistamine or  
 PT anti-inflammatory agents  
 XX Example 2; Fig 3; 44pp; English.  
 PS  
 XX This protein comprises tick Rhipicephalus appendiculatus (ra) novel  
 CC male-specific histamine binding protein 1 (MS-HBPI). Its amino  
 CC acid sequence was deduced from a cDNA clone (see AAV00227) obtained  
 CC from a salivary gland cDNA library. Female-specific HSP1 and  
 CC HSP2 (see AAM37446-47) and a related protein, D.KET6 (see AAM37449)  
 CC from Dermacentor reticulatus, were also identified. These novel  
 CC vasoactive amine binding proteins (VABPs) can be expressed in  
 CC host cells using e.g. a baculovirus expression system. They can  
 CC be used: (i) to assay histamine (or other VA such as serotonin) in  
 CC body fluids or cell culture supernatants, e.g. to monitor the  
 CC effect of allergens; (ii) for binding VA, e.g. to remove histamine  
 CC from blood, food, cell cultures etc.; (iii) as an antihistamine or  
 CC anti-inflammatory agents, e.g. for treating insect, snake or  
 CC scorpion bites or dermatitis; or as a carrier for slow release of  
 CC histamine-related compounds; (iv) in vaccines to protect against  
 CC metazoan parasites, especially in animals; (v) as reagents for the  
 CC studying inflammation, involvement of VA in ulcer formation for the  
 CC immune response etc. VABPs provide a more sensitive assay for  
 CC histamine than low-affinity antibodies currently used. They may  
 CC also be more effective and safer than conventional antihistamines.  
 XX Sequence 200 AA;  
 SQ  
 Query Match 28.1%; Score 267.5; DB 19; Length 200;  
 Best Local Similarity 39.3%; Pred. No. 2.1e-20;  
 Matches 70; Conservative 20; Mismatches 77; Indels 11; Gaps 7;  
 QY 3 PWADEAANGHODAMKHLQKLEBENDLIKATYKNDPVNGNDFTCVGTAQNLNDEKN 62  
 DB 20 PTWANEAKLGSYODAMKSLQDQDKRYLAQAQTGTVGVEEFTCVSAEKG--KKR 77  
 QY 63 VEAMFEMNADYVYQHTFEKATPDKMYGYNKENATYQTED--GOVLTVDLAFSD-DNC 119  
 DB 78 LNAFTILYKKNHLDKESHETITVWKAAYDTTENGKIKYETQGRTRQTFEDVFVSDYKNC 137  
 QY 120 YVITYALGPDGSGAG-YELMAT-DYTDVPASCLEKFNRYAA---GLPVDYTT-SDCLP 171  
 DB 138 DVLFVPERKSGSDGDEYELWVSEDKIDIPCCKFTMAFYAQOQEKYRNVYTDSSCKP 195  
 RESULT 12  
 AAY18080

ID AAY18080 standard; Protein; 200 AA.  
 AC AAY18080;  
 DT 06-AUG-1999 (first entry)  
 DE Histamine binding protein MS-HBP1.  
 KW Histamine binding protein; serotonin binding compound; inflammation;  
 KW gastric acid secretion; allergy; type I hypersensitivity reaction;  
 KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
 KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
 KW respiratory disease; coronary heart disease; cellular growth regulator;  
 KW tissue repair; blood-sucking ectoparasite; therapy.  
 OS Rhinipcephalus appendiculatus.  
 PN W09927104-A1.  
 PD 03-JUN-1999.  
 PF 26-NOV-1998; 98WO-GB03530.  
 PR 26-JUN-1998; 98GB-0013917.  
 PR 26-NOV-1997; 97GB-0025046.  
 PA (OXFO-) OXFORD VACS LTD.  
 PI Nuttall PA, Paesen GC;  
 DR WPI: 1999-357841/30.  
 DR N-PSDB: AAX76966.  
 PT Histamine and serotonin binding compounds useful for the treatment  
 PT of allergies  
 PS Claim 1; Fig 3; 84pp; English.  
 XX  
 XX This sequence is an example of a histamine or serotonin binding  
 CC compound (A), of the invention. The compounds are useful for regulating  
 CC the action of histamine and serotonin (in e.g. inflammation and gastric  
 CC acid secretion), the detection, quantification and removal of histamine  
 CC or serotonin (in animals, plants, cell cultures, food materials, or  
 CC humans) and in the treatment of various diseases and allergies,  
 CC (e.g. type I hypersensitivity reactions, urticaria, asthma, allergic  
 CC rhinitis (hay fever), atopic dermatitis, insect bites and food and drug  
 CC allergies, abnormal blood pressure, migraine, psychological disorders,  
 CC respiratory disease, and coronary heart disease). Histamine may also be  
 CC used to regulate cellular growth and tissue repair. The molecules may  
 CC also be used as components of vaccines directed against blood-sucking  
 CC ectoparasites.  
 CC  
 XX  
 SO Sequence 200 AA;  
 Query Match 28.1%; Score 267.5; DB 20; Length 200;  
 Best Local Similarity 39.3%; Pred. No. 2.1e-20;  
 Matches 70; Conservative 20; Mismatches 77; Indels 11; Gaps 7;  
 OY 3 PWADEANGSHODAMKHLQKLVENYDLIKATYKNDPWCNDFCTVCTAONINDEKN 62  
 DB 20 PTMANEAKISGYQDAMKSLQODONKRYLAQATOTDQWEEPTCVSVTAKEIG--KKK 77  
 OY 63 VEAFMFPMNNADYVQHTFEKATPDKMGYKKNENATYQTD--GQVLTDLAFSD--DNC 119  
 DB 78 LNAITILKNNKHLTLKESHERITVWKAQDYTTENGIKRTGTOTFEEDVFSDDKNC 137  
 OY 120 YVIALGPDGSGAG-YELMAT-DYTDVPASCLEKFNERYAA---GLVPADVTT-SDCLP 171  
 DB 138 DVIVFPERKSGDEBDYELMVSSEDKIDKIPDCCFTMAVFAQOQEKTYANVTDDSCRP 195

ID AAW37449 standard; Protein; 209 AA.  
 AC AAW37449;  
 DT 08-JUN-1998 (first entry)  
 DE Tick vasoactive amine binding protein D.RET6.  
 KW Vasoactive amine binding protein; D.RET6; histamine; serotonin;  
 KW assay; antihistamine; anti-inflammatory; insect bite; snake bite;  
 KW scorpion bite; dermatitis; vaccine; transgenic animal; tick.  
 OS Dermacenter reticularis.  
 PN W09744451-A2.  
 PD 27-NOV-1997.  
 PF 19-MAY-1997; 97WO-GB01372.  
 PR 18-APR-1997; 97GB-0007844.  
 PR 18-MAY-1996; 96GB-0010484.  
 PA (OXFO-) OXFORD VACS LTD.  
 PI Nuttall PA, Paesen GC;  
 DR WPI: 1998-018506/02.  
 DR N-PSDB: AAV00230.  
 PT New vasoactive amine binding proteins and related nucleic acid,  
 PT vectors - transformed cells and transgenic animals, used for  
 PT assaying or removing histamine and as antihistamine or  
 PT anti-inflammatory agents  
 PS Example 2; Fig 4; 44pp; English.  
 XX  
 XX This protein comprises tick Dermacenter reticularis (Dr) novel  
 CC vasoactive amine binding protein (VABP) D.RET6. Its amino acid  
 CC sequence was deduced from a cDNA clone (see AAV00230) obtained  
 CC from a Dr salivary gland cDNA library. 3 Novel VASPs, designated  
 CC FS-HBP1, FS-HBP2 and MS-HBP1 (see AAW37446-48), of the tick  
 CC Rhinipcephalus appendiculatus have also been identified. The  
 CC VABPs can be expressed in host cells using e.g. a baculovirus  
 CC expression system. They can be used: (i) to assay histamine (or  
 CC other VA such as serotonin) in body fluids or cell culture  
 CC supernatants, e.g. to monitor the effect of allergens; (ii) for  
 CC binding VA, e.g. to remove histamine from blood, food, cell  
 CC cultures etc.; (iii) as an antihistamine or anti-inflammatory  
 CC agents, e.g. for treating insect, snake or scorpion bites or  
 CC dermatitis, or as a carrier for slow release of histamine-related  
 CC compounds; (iv) in vaccines to protect against metazoan parasites,  
 CC especially in animals; (v) as reagents for studying inflammation,  
 CC involvement of VA in ulcer formation or the immune response etc.  
 CC VABPs provide a more sensitive assay for histamine than  
 CC low-affinity antibodies currently used. They may also be more  
 CC effective and safer than conventional antihistamines.  
 CC  
 XX  
 SO Sequence 209 AA;  
 Query Match 25.1%; Score 239; DB 19; Length 209;  
 Best Local Similarity 34.5%; Pred. No. 2.6e-17;  
 Matches 61; Conservative 24; Mismatches 82; Indels 10; Gaps 7;  
 OY 1 DKPYWADEANGSHODAMKHLQKLVENYDLIKATYKNDP-VGNDPFCVCTAONINDE 59  
 DB 29 ENPLMAHEELLGKRYODAMKSLDQGSVYVYLAKTYYENDGWSQFCLOVQOETERRRE 88  
 OY 60 EKNVEAFMFPMNNADTV-YQHTFEKATPDKMGY-KKENATYQTDQGVLTDLAFSD- 116

Db 89 DYTSTVETFRNASSPIKYYNVTETKAVFOYGYKKNIRNAIEYOVGGGLNITDILLPTDG 148  
QY 117 DNCVYIALGPDGSGAGYEIEMA--TDYTDVPASGLEFNEFYAA-GLFVRVYISDCL 170  
Db 149 ELCDVFYVPNAD--QGCELMVKKSHYKHVPDICTFEVNFVCAKDRRTYDIFNECV 202

## RESULT 14

AAV18081  
ID AAV18081 standard; Protein; 209 AA.

AAV18081;

06-AUG-1999 (first entry)

Histamine binding protein D.RET6.

KW Histamine binding protein; serotonin binding compound; inflammation;  
KW gastric acid secretion; allergy; type I hypersensitivity reaction;  
KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
KW respiratory disease; coronary heart disease; cellular growth regulator;  
KW tissue repair; blood-sucking ectoparasite; therapy.

Rhipicephalus appendiculatus.

MO9927104-A1.

03-JUN-1999.

26-NOV-1998; 98WO-GB03530.

26-JUN-1998; 98GB-0013917.

26-NOV-1997; 97GB-0025046.

(OXFO-) OXFORD VACS LTD.

Nuttall PA, Paesen GC;

WPI; 1999-357841/30.

N-PSDB; AAX76967.

Histamine and serotonin binding compounds useful for the treatment of allergies

Claim 1; Fig 4; 84pp; English.

This sequence is an example of a histamine or serotonin binding compound (A), of the invention. The compounds are useful for regulating the action of histamine and serotonin (in e.g. inflammation and gastric acid secretion), the detection, quantification and removal of histamine or serotonin (in animals, plants, cell cultures, food materials, or humans) and in the treatment of various diseases and allergies (e.g. type I hypersensitivity reactions, urticaria, asthma, allergic rhinitis (hay fever), atopic dermatitis, insect bites and food and drug allergies, abnormal blood pressure, migraine, psychological disorders, respiratory disease, and coronary heart disease). Histamine may also be used to regulate cellular growth and tissue repair. The molecules may also be used as components of vaccines directed against blood-sucking ectoparasites.

Sequence 209 AA;

Query Match 25.1%; Score 239; DB 20; Length 209;

Best Local Similarity 34.5%; Pred. No. 2.6e-17; Indels 10; Gaps 7;

Matches 61; Conservative 24; Mismatches 82;

QY 1 DKPVNADEANGEHODAMKHLQKLVENYDLIKATYKNDP-VGNDFTCVGTAQNINED 59  
Db 29 ENPLAMHEELGTYQDAMKSIDGVSVTYLAKTTEYENDTSGSGSFKCIQYGEIERKEE 88  
QY 60 EKNVEAMFMNANADTV-YQHTFEKATPDKMGY-NKENAIYQTEDEGOVLADVLAFSD- 116

Db 89 DYTSTVETFRNASSPIKYYNVTETKAVFOYGYKKNIRNAIEYOVGGGLNITDILLPTDG 148  
QY 117 DNCVYIALGPDGSGAGYEIEMA--TDYTDVPASGLEFNEFYAA-GLFVRVYISDCL 170  
Db 149 ELCDVFYVPNAD--QGCELMVKKSHYKHVPDICTFEVNFVCAKDRRTYDIFNECV 202

## RESULT 15

AAV18085  
ID AAV18085 standard; Protein; 203 AA.

AAV18085;

06-AUG-1999 (first entry)

Histamine binding protein Ih/Bm-HBP1.

KW Histamine binding protein; serotonin binding compound; inflammation;  
KW gastric acid secretion; allergy; type I hypersensitivity reaction;  
KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
KW respiratory disease; coronary heart disease; cellular growth regulator;  
KW tissue repair; blood-sucking ectoparasite; therapy.

Boophilus microplus.

Ixodes hexagonus.

Synthetic.

MO9927104-A1.

03-JUN-1999.

26-NOV-1998; 98WO-GB03530.

26-JUN-1998; 98GB-0013917.

26-NOV-1997; 97GB-0025046.

(OXFO-) OXFORD VACS LTD.

Nuttall PA, Paesen GC;

WPI; 1999-357841/30.

Histamine and serotonin binding compounds useful for the treatment of allergies

Claim 13; Fig 7; 84pp; English.

This sequence is an example of a histamine or serotonin binding compound (A), of the invention. cDNA encoding this sequence was isolated from a mixed Boophilus microplus/Ixodes hexagonus cDNA expression library. The compounds are useful for regulating the action of histamine and serotonin (in e.g. inflammation and gastric acid secretion), the detection, quantification and removal of histamine or serotonin (in animals, plants, cell cultures, food materials, or humans) and in the treatment of various diseases and allergies (e.g. type I hypersensitivity reactions, urticaria, asthma, allergic rhinitis (hay fever), atopic dermatitis, insect bites and food and drug allergies, abnormal blood pressure, migraine, psychological disorders, respiratory disease, and coronary heart disease). Histamine may also be used to regulate cellular growth and tissue repair. The molecules may also be used as components of vaccines directed against blood-sucking ectoparasites.

Sequence 203 AA;

Query Match 22.0%; Score 209.5; DB 20; Length 203;

Best Local Similarity 31.2%; Pred. No. 3.8e-14; Indels 7; Gaps 5;

Matches 50; Conservative 28; Mismatches 75;

QY 14 HODAMKHLQKLVENYDLIKATYKNDP-VGNDFTCVGTAQNINEDKENVAMFMNNA 73

Db	43	YODAKSINONVSTTYFFLSTYNNDSVWGKNFTCLSVTVTSKHSTFTVEYNTTKNOS	102
OY	74	DTVYQHFEKATPDKNMKGYNKENAITYQTEDEQVLTVDYLAESD-DNCYIYALGPDGSGA	132
Db	103	QQWVSMT-ENVTAVOEGYDVKNIIQMTTENNKKFNDTVVFTDGOCTDCLXIPYKEN--	158
OY	133	GYELMA-TDY-TDVPASCLEKFNENYAGLPVRDYYTSDCL	170
Db	159	GYELMVRSDYIQNTPTCCQFIEDLVALGRITYNISTPDCV	198

Search completed: August 1, 2003, 12:22:46  
Job time : 54.4019 secs

**THIS PAGE BLANK (USPTO)**





Db 79 KNEFAMFMNADTVQHFEKATPDKNMGYNENAITVOTEDGQVLTVDVLAESDNCY 138  
 QY 121 VYVAGPDGSGAGYEIMATDVTDPASCLEKFNENYAGLPVRYTSDCLPE 172  
 |||||  
 Db 139 VYVAGPDGSGAGYEIMATDVTDPASCLEKFNENYAGLPVRYTSDCLPE 190

RESULT 2  
 HBPM\_RHIAP STANDARD; PRT; 190 AA.  
 AC 077421;  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Female-specific histamine-binding protein 2 precursor (FS-HBP2).  
 OS Rhipicephalus appendiculatus (Brown ear tick).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodidae; Ixodidae; Rhipicephalus.  
 OX NCBI\_TaxID=34631;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND X-RAY CRYSTALLOGRAPHY (1.25 ANGSTROMS).  
 RC TISSUE=Salivary gland;  
 RA Paesen G.C., Adams P.L., Harlos K., Nuttall P.A., Stuart D.I.;  
 RT "Tick histamine-binding proteins: isolation, cloning, and three-dimensional structure.";  
 RL Mol. Cell 3:661-671(1999).  
 CC -1- FUNCTION: BINDS HISTAMINE WITH A HIGH-AFFINITY. THE ABILITY TO  
 OUTCOMPLETE HISTAMINE RECEPTORS INDICATES THAT ITS FUNCTION IS TO  
 CC SUPPRESS INFLAMMATION DURING BLOOD FEEDING.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U96081; AAC63107.1; -;  
 DR PDB; 1QFT; 19-APR-00.  
 DR InterPro; IPR002970; His\_binding.  
 DR Pfam; PF02098; His\_binding; 1.  
 DR ProDom; PD152455; His\_binding; 1.  
 DR Signal; 3D-structure.  
 FT CHAIN 1 19  
 FT SIGNAL 20 190  
 FT 2. FEMALE-SPECIFIC HISTAMINE-BINDING PROTEIN  
 FT DISULFID 67 188  
 FT DISULFID 138 167  
 FT TURN 23 24  
 FT HELIX 27 30  
 FT HELIX 31 33  
 FT HELIX 36 41  
 FT HELIX 42 45  
 FT STRAND 48 53  
 FT STRAND 57 58  
 FT TURN 59 61  
 FT STRAND 62 63  
 FT STRAND 64 64  
 FT TURN 66 76  
 FT STRAND 77 80  
 FT TURN 81 89  
 FT STRAND 91 92  
 FT TURN 97 106  
 FT STRAND 109 110  
 FT TURN 116 121  
 FT TURN 122 123  
 FT STRAND 126 135  
 FT TURN 136 137  
 FT STRAND 138 143

FT STRAND 152 157  
 FT TURN 160 161  
 FT HELIX 165 174  
 FT TURN 175 177  
 FT STRAND 181 182  
 FT TURN 186 188  
 SO SEQUENCE 190 AA; 21464 MW; 6923A3E902552B6F CRC64;

Query Match 66.2%; Score 630; DB 1; Length 190;  
 Best Local Similarity 66.5%; Pred. No. 1,3e-48;  
 Matches 113; Conservative 20; Mismatches 37; Indels 0; Gaps 0;

QY 1 DKPYWADEAANGEHODAMKHLQKLVENYDLIKATYVNDPVGNDFTCVGTAQNLNDE 60  
 :|||  
 Db 20 NOPMADEAANGAHQDAMKSLKADVENYVYKATYKNDPVGNDFTCVGVMAVDVNEDE 79  
 QY 61 KNEFAMFMNADTVQHFEKATPDKNMGYNENAITVOTEDGQVLTVDVLAESDNCY 120  
 |||||  
 Db 80 KSIQAEFLFMNADTVQHFEKATPDKNMGYNENAITVOTEDGQVLTVDVLAESDNCY 139

RESULT 3  
 HBPM\_RHIAP STANDARD; PRT; 200 AA.  
 AC 077422;  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DE Male-specific histamine-binding salivary precursor (MS-HBP).  
 OS Rhipicephalus appendiculatus (Brown ear tick).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodidae; Ixodidae; Rhipicephalus.  
 OX NCBI\_TaxID=34631;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Salivary gland;  
 RA Paesen G.C., Adams P.L., Harlos K., Nuttall P.A., Stuart D.I.;  
 RT "Tick histamine-binding proteins: isolation, cloning, and three-dimensional structure.";  
 RL Mol. Cell 3:661-671(1999).  
 CC -1- FUNCTION: BINDS HISTAMINE WITH A HIGH-AFFINITY. THE ABILITY TO  
 OUTCOMPLETE HISTAMINE RECEPTORS INDICATES THAT ITS FUNCTION IS TO  
 CC SUPPRESS INFLAMMATION DURING BLOOD FEEDING.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U96082; AAC63108.1; -;  
 DR HSSP; 077421; 1QFT.  
 DR InterPro; IPR002970; His\_binding.  
 DR Pfam; PF02098; His\_binding; 1.  
 DR ProDom; PD152455; His\_binding; 1.  
 DR Signal.  
 FT CHAIN 1 18  
 FT SIGNAL 19 200  
 FT 18. MALE-SPECIFIC HISTAMINE-BINDING SALIVARY  
 FT DISULFID 65 193  
 FT DISULFID 137 169  
 FT SEQUENCE 200 AA; 22851 MW; C46A1C8C6BCA008 CRC64;

Query Match 28.1%; Score 267.5; DB 1; Length 200;  
 Best Local Similarity 39.3%; Pred. No. 1.1e-16;

APPLICATION NUMBER: US/08/790,912  
FILING DATE: 29-JAN-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/026,939  
FILING DATE: 23-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Leary, Kathryn  
REGISTRATION NUMBER: 36,317  
REFERENCE/DOCKET NUMBER: 7600-401  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 567-2020  
TELEFAX: (215) 567-2991  
TELEX: 831-484  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2052 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-790-912-2

Query Match 8.5%; Score 81; DB 2; Length 2052;  
Best Local Similarity 24.1%; Pred. No. 15;  
Matches 45; Conservative 25; Mismatches 67; Indels 50; Gaps 8;

QY 12 GEHDDAKHLOKLVEN-----YDLKATYKNDPPWGNODTCVGTAAQNLNDEKNEA 65  
DB 1146 GTMLDASQIVSKAEINPLPTVEPLSTSGKD---SDFSKIAHYOANALVYKNEK 1201  
QY 66 WFMFMNNADYVYQHTFEKATPDKMVGYKNEAITYOTE-----DGOVLTDLAFSD 117  
DB 1202 LLRPYN-----KSTIVKGNLVKENSILYKLELSAVMMKDDQVTDIVS-NKO 1249  
QY 118 NCYVIVALGPDGSGAGYEL-WATDYTDVPASCLEKFEYAG-----LPVRDYY 165  
DB 1250 TANKLLHYNDHSEKFDLKYQTDFAANLP-----EYNIGNTGLYTPNQFLYDRDSI 1301  
QY 166 TSDCLPE 172  
DB 1302 VKEVLPE 1308

RESULT 5  
US-08-271-354-9

Sequence 9, Application US/08271354  
Patent No. 5695983  
GENERAL INFORMATION:  
APPLICANT: Miller, Samuel I.  
APPLICANT: Melanos, John J.  
TITLE OF INVENTION: SALMONELLA VACCINES  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/271,354  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/077,526  
FILING DATE: July 9, 1993  
APPLICATION NUMBER: 07/629,602

FILING DATE: December 18, 1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/220001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 246 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-271-354-9

Query Match 8.4%; Score 80; DB 1; Length 246;  
Best Local Similarity 23.7%; Pred. No. 0.84;  
Matches 27; Conservative 21; Mismatches 46; Indels 20; Gaps 5;

QY 38 NDPVWGN---DFTCVGTAAQNLNDEKNEAWFMFMNNADTV-----YQHTFEKAT 85  
DB 110 NNGYWGSPVNLTFGTITEE-OKOELKALFKF---DSINFSIIPERIOETIRAN 164  
QY 86 PDKMYGYNKENAITYOTE---DGOVLTDLAFSDNCRVIVALGPDGSGAGYEL 136  
DB 165 ASGIISVTESDIVRAEIHNGSEYDITITANTARAVMTLNKDSIAGEYI 218

RESULT 6

US-08-565-861-9

Sequence 9, Application US/08565861  
Patent No. 5843426  
GENERAL INFORMATION:  
APPLICANT: Miller, Samuel I.  
APPLICANT: Melanos, John J.  
APPLICANT: Hohmann, Elizabeth  
TITLE OF INVENTION: SALMONELLA VACCINES  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/565,861  
FILING DATE: 01-DEC-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/271,354  
FILING DATE: 06-JUL-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/090,526  
FILING DATE: 09-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/629,602  
FILING DATE: 18-DEC-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/303001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
TELEX: 200154



OY 122 IYALGPDGSGAGYELMATDYTPVASCLEKENEYAAGLPVADYVT 166  
Db 91 -----PDIP-----QLMAOKITALNYSGL-NLNQGVASPCFTICT 125

## RESULT 13

hypothetical protein C37A5.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jul-2000  
C:Accession: J19806  
R:White, S.  
Submitted to the EMBL Data Library, March 1997  
A:Accession: J19806  
A:Reference number: Z19180  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-512 <WIL>  
A:Cross-references: EMBL:Z92828; PIDN:CA807337.1; GSPDB:GN00019; CESP:C37A5.1  
A:Experimental source: clone C37A5  
C:Genetics:  
A:Gene: CESP:C37A5.1  
A:Map position: 1  
A:Introns: 19/3; 60/2; 162/1; 265/3; 313/3; 364/2; 388/1; 467/3  
C:Superfamily: Caenorhabditis elegans hypothetical protein C01B12.5

## Query Match

Best Local Similarity 8.6%; Score 81.5; DB 2; Length 512;  
Matches 31; Conservative 17; Mismatches 46; Indels 23; Gaps 5;

OY 46 FTTCVGTAAQNLNEDKNEAFMFMMNNADTYVQHTFEKATPDKMYGYNKRNATYQTEEDG 105  
Db 249 FTCLFTQHLDMEPTKTIDYFPL-----TVEQTF-----FMGMKV-----A 288  
OY 106 QVLTDLVAFSDNCYVYIALGPDGSGAGYELMATDYTPVASCLEKENE--YAAGLP 160  
Db 289 EALNPLGDDDDDECNYLIDRN-IATGMATVNSKYSVPMLADKFNPSYADYYP 344

## RESULT 14

S28179  
alpha-amylase (EC 3.2.1.1) - Streptomyces lividans  
C:Species: Streptomyces lividans  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 24-Apr-1998  
C:Accession: S28179  
R:Tsero, L.S.; Lin, L.L.; Chen, J.C.; Chen, J.H.; Hsu, W.H.  
Biochim. Biophys. Acta 1171, 255-262, 1993  
A:Title: Cloning and characterization of an alpha-amylase gene from Streptomyces lividans  
A:Reference number: S28179; MUID:93144336; PMID:8424949  
A:Accession: S28179  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-919 <TSA>  
C:Genetics:  
A:Start codon: GTG  
C:Keywords: glycosidase; hydrolase

Query Match 8.6%; Score 81.5; DB 2; Length 919;  
Best Local Similarity 27.0%; Pred. No. 31;  
Matches 34; Conservative 11; Mismatches 32; Indels 49; Gaps 7;

OY 23 KLVEENTDLIKA-----TYKNDPV-----WGNDFTCV-----GTAQAQ 54  
Db 101 KGLTEKLDYIKGLGTTSTMAPIFKNQPVQGTGDASAGYGYWTDTPTQVDPRHGT--- 157  
OY 55 NLNDEKN-----VEAMFMNNADTYVQHTFEKAT-PDKMYGYNKRNATYQTEED 104  
Db 158 --NRDLKTLISKAAKMKVFE-----DVITNHTADVVDYEKESYDLSKGAFFYLTKD 209  
OY 105 GQVLTLD 110  
Db 210 GQPFDD 215

## RESULT 15

hypothetical protein MYPV\_2120 [Imported] - Mycoplasma pulmonis (strain UAB CTIP)  
D90538  
C:Species: Mycoplasma pulmonis  
C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001  
C:Accession: D90538  
R:Chambaud, I.; Hellig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galissou, F.; Moszer, Nucleic Acids Res. 29, 2145-2153, 2001  
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma p  
A:Reference number: A99512; MUID:21267165; PMID:11353084  
A:Accession: D90538  
A:Status: preliminary  
A:Accession: D90538  
A:Molecule type: DNA  
A:Residues: 1-1377 <KUR>  
A:Cross-references: GB:A445566; PID:914089625; PIDN:CA613385.1; GSPDB:GN00153  
A:Experimental source: strain UAB CTIP  
C:Genetics:  
A:Gene: MYPV\_2120  
A:Genetic code: SGC3

## Query Match

Best Local Similarity 8.6%; Score 81.5; DB 2; Length 1377;  
Matches 41; Conservative 19; Mismatches 61; Indels 79; Gaps 8;

OY 13 EHODAMK-----HLQKLVENYDL-----IKATYK 37  
Db 1114 KRQDAWSEDEKYYVDQKANKKTIASQYKLDQYFKIRKNSKDALKEITPLGTTINLTIN 1173  
OY 38 NDPYWGNDFTCVGTAAQNLNEDKNEAFMFMMNNADTYVQHTFEKATPDKMYGYNKEN- 96  
Db 1174 EDKTVRADFA-----DKKEIDKIKRYK--DLEIQSFE-----GVTNSNG 1212  
OY 97 -----AITYQTEEDQVLTIVLAFSDNCYVYIALGPDGSGAGYELMATDYTPVASCLEKF 152  
Db 1213 EQIVRNQNEEAHAVERKNLLSSDNNYSKYSLSVD-----QINAFKKKILVREARKF 1266  
OY 153 NEYAAGLPVRDVTSDCLPE 172  
Db 1267 S-----NSDLFPD 1274

Search completed: August 1, 2003, 12:16:32  
Job time : 15.7771 secs

**THIS PAGE BLANK (USPTO)**

Db 288 AFTVLA-MMDYPTDFELGFLPANPVKCGDRLLSEARITGLRALAGLVNAGSEHCY 346  
OY 121 VITAL-----GPDGSGAGYELMATDYTDVPASCLEKFNEXAAGLPVRDY 165  
Db 347 DIRLRLHSCADPTGCGTGPBARMDY---QACTE-INLTFASNVTDMF 391

## RESULT 13

US-09-388-413B-6  
; Sequence 6, Application US/09388413B  
; Patent No. 6485955  
; GENERAL INFORMATION:  
; APPLICANT: Huber, Brigitte T.  
; TITLE OF INVENTION: QUISCENT CELL DIPEPTIDYL PEPTIDASE: A NOVEL CYTOPLASMIC SERINE  
; FILE REFERENCE: 50420 CIP  
; CURRENT APPLICATION NUMBER: US/09/388,413B  
; PRIOR FILING DATE: 1999-09-01  
; PRIOR FILING DATE: 1997-10-06  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 6  
; LENGTH: 537  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-388-413B-6

Query Match 8.4% Score 79.5; DB 4; Length 537;  
Best Local Similarity 25.9%; Pred. No. 3;  
Matches 44; Conservative 27; Mismatches 58; Indels 41; Gaps 12;

OY 15 QDAMKHLQKLVEENYDLIKATYKNDPVGNDFTCVGTAQNLNDEKNEVEMFMNN 72  
Db 221 REAFROKDLFLGAGDVT-----W--EFGTC-----QPLS-DEKDTQLFMRARN 264  
OY 73 ADVYQHTFEKATPDKMYGYNKENAITYYQTE---DQGVLTDLVAFS-----DDNCY 120  
Db 265 AFTVLA-MMDYPTDFELGFLPANPVKCGDRLLSEARITGLRALAGLVNAGSEHCY 323  
OY 121 VITAL-----GPDGSGAGYELMATDYTDVPASCLEKFNEXAAGLPVRDY 165  
Db 324 DIRLRLHSCADPTGCGTGPBARMDY---QACTE-INLTFASNVTDMF 368

## RESULT 14

US-08-284-941-2  
; Sequence 2, Application US/08284941  
; Patent No. 5863756  
; GENERAL INFORMATION:  
; APPLICANT: BARR, PHILIP J  
; APPLICANT: KIEFER, MICHAEL C  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PACE 4 AND  
; TITLE OF INVENTION: PACE 4.1 GENE AND POLYPEPTIDES IN CELLS  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: COOLEY GODNARD CASTRO HUDDLESON & TATUM  
; STREET: FIVE PALO ALTO SQUARE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/284,941  
; FILING DATE: 2 August 1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: NEELEY PH.D., RICHARD L.

; REGISTRATION NUMBER: 30092  
; REFERENCE/DOCKET NUMBER: CHIR-009/01US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 843-5070  
; TELEFAX: (415) 857-0663  
; TELEFAX: 380816 COOLEY PA  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 969 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-284-941-2

Query Match 8.1% Score 77.5; DB 2; Length 969;  
Best Local Similarity 21.3%; Pred. No. 12;  
Matches 39; Conservative 28; Mismatches 81; Indels 35; Gaps 7;

OY 13 EHODAMKHLQKLVEENYDLIKATYKNDPVGNDFTCVGTAQNLNDEKNEVEMFMNN 68  
Db 139 QDQEVRRVRRKQYRSD---PQALYFNDPIISMNMYLHCGDKNSRCRENNVQAMKRGYT 195  
OY 69 FMNNADTVYQHTFEKATPDKMYGYNK-----ENAITY-----QTEGCVLTDLVLA 113  
Db 196 GKNVYVYIIDDGIERNHPDLAPNYDSYASYVNGNDYDPSPRYDASNNHGRCAGEVA 255  
OY 114 FSDQNCYVIALGPDGSGAGYELMATDYTDVPASCLEKFNEXAAGLPVR---DYVNSDC 169  
Db 256 ASANNSTCYIGVIAVNAKIGIRLMDGDVTVVE-----AKSLGIRPNYIDYISASW 306  
OY 170 LPE 172  
Db 307 GPD 309

## RESULT 15

US-08-447-642-2  
; Sequence 2, Application US/08447642  
; Patent No. 5989890  
; GENERAL INFORMATION:  
; APPLICANT: BARR, PHILIP J  
; APPLICANT: KIEFER, MICHAEL C  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PACE 4 AND  
; TITLE OF INVENTION: PACE 4.1 GENE AND POLYPEPTIDES IN CELLS  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: COOLEY GODNARD CASTRO HUDDLESON & TATUM  
; STREET: FIVE PALO ALTO SQUARE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/447,642  
; FILING DATE: 23-MAY-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/284,941  
; FILING DATE: 2 August 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: NEELEY PH.D., RICHARD L.  
; REGISTRATION/DOCKET NUMBER: 30092  
; REFERENCE/DOCKET NUMBER: CHIR-009/01US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 843-5070  
; TELEFAX: (415) 857-0663  
; TELEFAX: 380816 COOLEY PA  
; INFORMATION FOR SEQ ID NO: 2:

```

; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/090,112
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: 60/090,113
; EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 518
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-461-325-518

Query Match
Best Local Similarity 25.9%; Pred. No. 2.7;
Matches 44; Conservative 27; Mismatches 58; Indels 41; Gaps 12;

8.4%; Score 79.5; DB 4; Length 490;

OY 15 ODAWKHLQKL-VEENYDLIKATYKNDPVWGNDP-TCVGTAAQNLNEDKNEVAFMFMNN 72
DB 219 REAFRQIKDLFLQGAAYTVR-----W-EFGTC-----QPLS-DEKDLTQLFMFRN 262
OY 73 ADYVQHTFEKATPDKMGYKKNATYYQTE---DGOVLTDVLAFS-----DDNCY 120
DB 263 AFTVLA-MMDYPTDFELGPLPAPNVKVCRLSEAGRTIGLRALAGLVYNASGSEHCY 321
OY 121 VIVAL-----GPDGSGAGYELMATDYTDVPASCLEKFEVYAAGLPVRDYY 165
DB 322 DIYRLYHSCADPTGCGTGPDARANDY---QACTE-INLTFASNNVTDMF 366

RESULT 10
US-09-794-236-2
; Sequence 2, Application US/09794236
; Patent No. 6337069
; GENERAL INFORMATION:
; APPLICANT: Grouzmann, Eric
; APPLICANT: Lacroix, Jean-Silvain
; APPLICANT: Monod, Michel
; TITLE OF INVENTION: Method of Treating Rhinitis and Sinusitis
; FILE REFERENCE: 81985/276823
; CURRENT APPLICATION NUMBER: US/09/794,236
; CURRENT FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 2
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-236-2

Query Match
Best Local Similarity 25.9%; Pred. No. 2.7;
Matches 44; Conservative 27; Mismatches 58; Indels 41; Gaps 12;

8.4%; Score 79.5; DB 4; Length 492;

OY 15 ODAWKHLQKL-VEENYDLIKATYKNDPVWGNDP-TCVGTAAQNLNEDKNEVAFMFMNN 72
DB 221 REAFRQIKDLFLQGAAYTVR-----W-EFGTC-----QPLS-DEKDLTQLFMFRN 264
OY 73 ADYVQHTFEKATPDKMGYKKNATYYQTE---DGOVLTDVLAFS-----DDNCY 120
DB 265 AFTVLA-MMDYPTDFELGPLPAPNVKVCRLSEAGRTIGLRALAGLVYNASGSEHCY 323
OY 121 VIVAL-----GPDGSGAGYELMATDYTDVPASCLEKFEVYAAGLPVRDYY 165
DB 324 DIYRLYHSCADPTGCGTGPDARANDY---QACTE-INLTFASNNVTDMF 368

RESULT 11
US-09-345-469-1
; Sequence 1, Application US/09345469
; Patent No. 6369210
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
```

```

; APPLICANT: Macbeth, Kyle J.
; APPLICANT: Williamson, Mark
; TITLE OF INVENTION: 22012, A No. 6369210e1 Human Carboxypeptidase
; FILE REFERENCE: 5800-38
; CURRENT APPLICATION NUMBER: US/09/345,469
; CURRENT FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-345-469-1

Query Match
Best Local Similarity 25.9%; Pred. No. 2.7;
Matches 44; Conservative 27; Mismatches 58; Indels 41; Gaps 12;

8.4%; Score 79.5; DB 4; Length 492;

OY 15 ODAWKHLQKL-VEENYDLIKATYKNDPVWGNDP-TCVGTAAQNLNEDKNEVAFMFMNN 72
DB 221 REAFRQIKDLFLQGAAYTVR-----W-EFGTC-----QPLS-DEKDLTQLFMFRN 264
OY 73 ADYVQHTFEKATPDKMGYKKNATYYQTE---DGOVLTDVLAFS-----DDNCY 120
DB 265 AFTVLA-MMDYPTDFELGPLPAPNVKVCRLSEAGRTIGLRALAGLVYNASGSEHCY 323
OY 121 VIVAL-----GPDGSGAGYELMATDYTDVPASCLEKFEVYAAGLPVRDYY 165
DB 324 DIYRLYHSCADPTGCGTGPDARANDY---QACTE-INLTFASNNVTDMF 368

RESULT 12
US-09-461-325-219
; Sequence 219, Application US/09461325A
; Patent No. 6475753
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/09/461,325A
; CURRENT FILING DATE: 1999-12-14
; EARLIER APPLICATION NUMBER: PCT/US99/13418
; EARLIER FILING DATE: 1999-06-15
; EARLIER APPLICATION NUMBER: 60/089,507
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,508
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,509
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,510
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/090,112
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: 60/090,113
; EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 219
; LENGTH: 515
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-461-325-219

Query Match
Best Local Similarity 25.9%; Pred. No. 2.9;
Matches 44; Conservative 27; Mismatches 58; Indels 41; Gaps 12;

8.4%; Score 79.5; DB 4; Length 515;

OY 15 ODAWKHLQKL-VEENYDLIKATYKNDPVWGNDP-TCVGTAAQNLNEDKNEVAFMFMNN 72
DB 244 REAFRQIKDLFLQGAAYTVR-----W-EFGTC-----QPLS-DEKDLTQLFMFRN 287
OY 73 ADYVQHTFEKATPDKMGYKKNATYYQTE---DGOVLTDVLAFS-----DDNCY 120
```



SEQUENCE CHARACTERISTICS:  
 LENGTH: 969 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-447-642-2

Query Match. 8.1%; Score 77.5; DB 2; Length 969;  
 Best Local Similarity 21.3%; Pred. No. 12; Mismatches 7;  
 Matches 39; Conservative 28; Indels 35; Gaps 7;

QY	13	EHODAMKHLQKLV	EEYNDLIKATYKNDP	VGNDFTCVGTAAONLNEDEKNVE-AW---	FM 68
DB	139	QOQEVKRRVKKQVRSD---	PGALYFNDPIWNSMNYLHCGDKNSRCRSEMANVQAAMKRGYT		195
QY	69	FMNNADIVYQHTFEKATPDKMKGYNK----	ENAITY-----	QTEGQVLTDVLA	113
DB	196	GKNVYVYTLDDGIERNHDLAPNYDSYASYDVNGNDYDPSPRYDASNENKHGTRCAGEVA			255
QY	114	FSDDMCYIYIALGPDGSGAGYELMTDYPDVPASCLEKFN	EYAAGLPVR---	DVYTSDC	169
DB	256	ASANSSTCIYGVIAVNAKIGIRMLDGDVTDVVE-----	AKSLGIRPNYIDLYSASW		306
QY	170	LPE 172			
DB	307	GPD 309			

Search completed: August 1, 2003, 12:17:19  
 Job time: 16.1048 secs

**THIS PAGE BLANK (USPTO)**

ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.  
STREET: 1601 Market Street, 36th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103-2398

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/790,912  
FILING DATE: 29-JAN-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/026,939  
FILING DATE: 23-SEP-1996

ATTORNEY/AGENT INFORMATION:  
NAME: Leary, Kathryn  
REGISTRATION NUMBER: 36,317  
REFERENCE/DOCKET NUMBER: 7600-401  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 567-2020  
TELEFAX: (215) 567-2991  
TELEX: 831-494

INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1861 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-790-912-4

Query Match 8.68; Score 82; DB 2; Length 1861;  
Best Local Similarity 27.18; Pred. No. 10;  
Matches 51; Conservative 12; Mismatches 47; Indels 78; Gaps 12;

QY 7 DEANAGEHODAMKHLKLVEN-----YD-----LIKATYK 37  
DB 1370 DRRNNEHR-AGELMKFVEDNAOETAKKROHDYDWRILDSOGREKRLNLYDAYK 1428  
QY 38 NDVWGNDFCTGVTAAQNLNEDENYEAWEFMNNA-----DYVYQHTFEKATPD 87  
DB 1429 -----FDDDTYV-----DKATVEAQFDSNPAMKYPFGVGNKKVHNKRGAVATGD 1474  
QY 88 KMY--GY---NKENATY---QTEDEQVLTDLVAFSDDNKYVYALGPDG--SGAGYELW 137  
DB 1475 SVYKMGYRMLDKDGAITYTHEMTHD-----SDNELY---LEGYGRSRGLGPEFF 1520

QY 138 ATDYTDVP 145  
DB 1521 AKGLQAP 1528

RESULT 3  
US-08-790-912-3  
Sequence 3, Application US/08790912  
Patent No. 5976542  
GENERAL INFORMATION:  
APPLICANT: Weiser, Jeffrey N.  
APPLICANT: Plaut, Andrew G.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT  
TITLE OF INVENTION: OF STREPTOCOCCUS PNEUMONIAE INFECTION  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.  
STREET: 1601 Market Street, 36th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103-2398

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/790,912  
FILING DATE: 29-JAN-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/026,939  
FILING DATE: 23-SEP-1996

ATTORNEY/AGENT INFORMATION:  
NAME: Leary, Kathryn  
REGISTRATION NUMBER: 36,317  
REFERENCE/DOCKET NUMBER: 7600-401  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 567-2020  
TELEFAX: (215) 567-2991  
TELEX: 831-494

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1964 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-790-912-3

Query Match 8.58; Score 81; DB 2; Length 1964;  
Best Local Similarity 24.18; Pred. No. 14;  
Matches 45; Conservative 25; Mismatches 67; Indels 50; Gaps 8;

QY 12 GEHODAMKHLKLVEN-----YDLIKATYKNDPVWGNDFCTGVTAAQNLNEDENVEA 65  
DB 1073 GIMLDSQYSKAEINPLTLPTVEPLSTSGKD-----SDFSKIANYQARALVYKNIEK 1128  
QY 66 WFMFNNAADYVYQHTFEKATPDKMYGNKENATYQTE-----DGOVLTDLVAFSD 117  
DB 1129 LLPFYV-----KSTIVYGNLVKENSLLYQKELLSAVMMKMDQVITDYS-NMQ 1176  
QY 118 NCYVYIALGPDGAGCYEL-WATDYTDVPASCLEREKNEVNAAG-----LPVRDY 165  
DB 1177 TANKLLHYNDHSEKEDLKYPQDFANLP-----EYNLGMTGLYTPNDFLYDRDSI 1228  
QY 166 TSDCLPE 172  
DB 1229 VKEVLP 1235

RESULT 4  
US-08-790-912-2  
Sequence 2, Application US/08790912  
Patent No. 5976542  
GENERAL INFORMATION:  
APPLICANT: Weiser, Jeffrey N.  
APPLICANT: Plaut, Andrew G.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT  
TITLE OF INVENTION: OF STREPTOCOCCUS PNEUMONIAE INFECTION  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.  
STREET: 1601 Market Street, 36th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103-2398

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2003, 12:15:04 ; Search time 13.1048 Seconds  
(without alignments)  
555.330 Million cell updates/sec

Title: US-10-087-195-6  
Perfect score: 952  
Sequence: 1 DKPYMADEANGHEQDAMKH.....NEYAAGLPVRDYTSDCIPE 172

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Issued\_Patents\_AA.\*  
1: /cgn2-6/ptodata/1/1aa/5A-COMB.pep.\*  
2: /cgn2-6/ptodata/1/1aa/5B-COMB.pep.\*  
3: /cgn2-6/ptodata/1/1aa/6A-COMB.pep.\*  
4: /cgn2-6/ptodata/1/1aa/6B-COMB.pep.\*  
5: /cgn2-6/ptodata/1/1aa/PCTUS-COMB.pep.\*  
6: /cgn2-6/ptodata/1/1aa/Backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	84.5	8.9	469	US-09-131-750-29	Sequence 29, Appl
2	82	8.6	1861	US-08-790-912-4	Sequence 4, Appl
3	81	8.5	1964	US-08-790-912-3	Sequence 3, Appl
4	81	8.5	2052	US-08-790-912-2	Sequence 2, Appl
5	80	8.4	246	US-08-271-354-9	Sequence 9, Appl
6	80	8.4	246	US-08-565-861-9	Sequence 9, Appl
7	80	8.4	246	PCT-US94-07653-9	Sequence 9, Appl
8	79.5	8.4	490	US-09-461-325-250	Sequence 250, App
9	79.5	8.4	492	US-09-461-325-518	Sequence 518, App
10	79.5	8.4	492	US-09-794-236-2	Sequence 2, Appl
11	79.5	8.4	492	US-09-345-469-1	Sequence 1, Appl
12	79.5	8.4	515	US-09-461-325-219	Sequence 2, Appl
13	79.5	8.4	537	US-09-388-413B-6	Sequence 6, Appl
14	77.5	8.1	969	US-08-284-941-2	Sequence 2, Appl
15	77.5	8.1	969	US-08-447-642-2	Sequence 2, Appl
16	77.5	8.1	969	US-09-236-503-2	Sequence 2, Appl
17	77.5	8.1	969	PCT-US93-02147A-2	Sequence 2, Appl
18	76.5	8.0	550	US-08-107-532A-6382	Sequence 6382, Ap
19	76.5	8.0	958	US-08-426-236-4	Sequence 4, Appl
20	75.5	7.9	187	US-07-630-551B-3	Sequence 3, Appl
21	75.5	7.9	461	PCT-US93-12464-2	Sequence 2, Appl
22	75.5	7.9	617	US-09-252-991A-27584	Sequence 27584, A
23	74.5	7.8	459	US-08-220-151-12	Sequence 12, Appl
24	74.5	7.8	459	US-08-220-151-14	Sequence 14, Appl
25	74.5	7.8	459	US-08-413-118-12	Sequence 12, Appl
26	74.5	7.8	459	US-08-413-118-14	Sequence 14, Appl
27	74.5	7.8	459	US-08-473-446-12	Sequence 12, Appl

28	74.5	7.8	459	US-08-473-446-14	Sequence 14, Appl
29	74.5	7.8	459	US-09-213-053-6	Sequence 6, Appl
30	74.5	7.8	1394	US-09-213-053-2	Sequence 2, Appl
31	74	7.8	547	US-08-083-948-8	Sequence 8, Appl
32	74	7.8	547	US-08-393-785-8	Sequence 8, Appl
33	74	7.8	547	US-08-475-694-8	Sequence 8, Appl
34	74	7.8	547	US-08-712-057-8	Sequence 8, Appl
35	74	7.8	547	US-09-347-878-36	Sequence 36, Appl
36	73.5	7.7	337	US-09-134-001C-3207	Sequence 3207, Ap
37	73.5	7.7	881	US-08-525-940-21	Sequence 21, Appl
38	73.5	7.7	881	US-08-976-838-21	Sequence 21, Appl
39	73.5	7.7	915	US-08-525-940-18	Sequence 18, Appl
40	73.5	7.7	915	US-08-976-838-18	Sequence 18, Appl
41	73.5	7.7	915	US-09-214-5558-7	Sequence 7, Appl
42	73	7.7	182	US-08-828-741B-2	Sequence 2, Appl
43	73	7.7	182	US-09-160-567-2	Sequence 2, Appl
44	73	7.7	182	US-09-710-289-2	Sequence 2, Appl
45	73	7.7	2037	US-09-306-998-3	Sequence 3, Appl

#### ALIGNMENTS

```
RESULT 1
US-09-131-750-29
: Sequence 29, Application US/09131750A
: Patent No. 6361964
: GENERAL INFORMATION:
: APPLICANT: Kaiser, Chris A.
: TITLE OF INVENTION: EUKARYOTIC DISULFIDE BOND-FORMING
: FILE REFERENCE: 01997/502002
: CURRENT APPLICATION NUMBER: US/09/131,750A
: EARLIER FILING DATE: 1998-08-10
: EARLIER APPLICATION NUMBER: 60/055,586
: NUMBER OF SEQ ID NOS: 29
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 29
: LENGTH: 469
: TYPE: PRT
: ORGANISM: Homo sapiens, Mus musculus, Rattus norvegicus
US-09-131-750-29

Query Match      8.9%; Score 84.5; DB 4; Length 469;
Best Local Similarity 25.3%; Pred. No. 0.69; 80; Indels 21; Gaps 8;
Matches 44; Conservative 29; Mismatches 15

QY      3 PVMADENANGHEQDAMKHLQKLVENYDLIK-ATYKNDPVWGN-----DFTCVGTAON 55
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      86 PIMWDISQCGRRDCAVPCQS--DEVPDGKISAKYKSEENNLIEECQAEKRGANDES 143

QY      56 LNEED-EKNVEAMFEMNNADTVYQHTFEKATPDKMYG---YKNENATYOTEDGOVLTV 111
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      144 LSEETQAVLQWTHDDSDNFCADDIOSEAEVYDLLNPERYGYGKPDAMKIMNV 203

QY      112 LAESDDNVCY-----VITVLPDGSAGYEELMATDYTPASCLE--FNEYAAGL 159
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      204 I-YEENCFFKOTIKRPLNPLASGQTSSENTFYSMLEGLCVERAFRYRLISGL 255

RESULT 2
US-08-790-912-4
: Sequence 4, Application US/08790912
: Patent No. 5976542
: GENERAL INFORMATION:
: APPLICANT: Weiser, Jeffrey N.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
```

SOFTWARE: Patentin Ver. 2.0  
 SEQ ID NO 14  
 LENGTH: 481  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-989-545-14

Query Match 8.1%; Score 77; DB 10; Length 481;  
 Best Local Similarity 21.3%; Pred. No. 6.1;  
 Matches 40; Conservative 33; Mismatches 63; Indels 52; Gaps 9;

OY 2 KPWADDEAANGEHODANKHLOK---LVEENYDLIKATYKNDPVWGNDFTC-----VGTAA 53  
 Db 102 KPCQSDPEVPDGIKSASYKSEANNLIEE-----CEQAEKLGAVD 141  
 OY 54 QNLNED-EKNVEAFMFEMNNDVTYQ-----HFEKATPDKMTG---YKNENA 97  
 Db 142 ESLSSEETOKAVLQMTKHDSSDNFCCEADNMCWSSKKKVFDIQSPAEYVDLLNPERY 201  
 OY 98 ITVQTEDEGOVLTDLAFSDNDCY---VIVALGPDGSGAGYELWATDYTDVPAACLEK-- 151  
 Db 202 TGYKGPDAKIMNVI--YEENCFRPTIKRPLNPLASGOGTSEENTFTSWLEGLCYEKRA 259  
 OY 152 FNEYAAGL 159  
 Db 260 FYRLISGL 267

Search completed: August 1, 2003, 12:23:44  
 Job time : 18.0362 secs



Db 263 AFTVLA-MMDYPTFLGLPAPNPKVCGCDRLSEAOQRTTGLRALAGLVNAGSGSEHCY 321  
OY 121 VIVAL-----GPDGSGAGYELMATDVTDPASCLEKFNENYAAGLPVRDY 165  
Db 322 DLYRLHSCADPTGCGTGPDPARAMDY---QACTE-INLTFASNNTDMF 366

RESULT 9  
US-10-068-134-1  
; Sequence 1, Application US/10068134  
; Publication No. US20020156264A1  
; GENERAL INFORMATION:  
; APPLICANT: Kapellier-Libermann, Rosana  
; APPLICANT: Macbeth, Kyle J.  
; APPLICANT: Williamson, Mark  
; TITLE OF INVENTION: 22012, A No. US20020156264A1el Human Carboxypeptidase  
; FILE REFERENCE: 5800-38  
; CURRENT APPLICATION NUMBER: US/10/068,134  
; CURRENT FILING DATE: 2002-02-06  
; PRIOR APPLICATION NUMBER: US/09/345,469  
; PRIOR FILING DATE: 1999-06-30  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 492  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-068-134-1

Query Match  
Best Local Similarity 25.9%; Score 79.5; DB 14; Length 492;  
Best Local Similarity 25.9%; Pred. No. 3.3;  
Matches 44; Conservative 27; Mismatches 58; Indels 41; Gaps 12;

OY 15 ODAWKHLQKL-VEENTDLIKATYKNDPVWGNDP-TGCVGTAAQNLNDEKKNVEMFPMNN 72  
Db 221 REAFROIKDLFLOGAIDTVR-----W-EFGTC-----OPLS-DEKDLTQLFPMFARN 264  
OY 73 ADVYQHTFEKATPDKMYGKKNENAITTYOTE---DGOVLTDLVAFS-----DDNCCY 120  
Db 265 AFTVLA-MMDYPTFLGLPAPNPKVCGCDRLSEAOQRTTGLRALAGLVNAGSGSEHCY 323  
OY 121 VIVAL-----GPDGSGAGYELMATDVTDPASCLEKFNENYAAGLPVRDY 165  
Db 324 DLYRLHSCADPTGCGTGPDPARAMDY---QACTE-INLTFASNNTDMF 368

RESULT 10  
US-10-012-542-219  
; Sequence 219, Application US/10012542  
; Publication No. US20030044851A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 94 Human Secreted Proteins  
; FILE REFERENCE: P2029P1  
; CURRENT APPLICATION NUMBER: US/10/012,542  
; CURRENT FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22  
; NUMBER OF SEQ ID NOS: 532  
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 219  
; LENGTH: 515  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-012-542-219

Query Match  
Best Local Similarity 25.9%; Score 79.5; DB 15; Length 515;  
Best Local Similarity 25.9%; Pred. No. 3.6;  
Matches 44; Conservative 27; Mismatches 58; Indels 41; Gaps 12;

OY 15 ODAWKHLQKL-VEENTDLIKATYKNDPVWGNDP-TGCVGTAAQNLNDEKKNVEMFPMNN 72  
Db 244 REAFROIKDLFLOGAIDTVR-----W-EFGTC-----OPLS-DEKDLTQLFPMFARN 287  
OY 73 ADVYQHTFEKATPDKMYGKKNENAITTYOTE---DGOVLTDLVAFS-----DDNCCY 120  
Db 288 AFTVLA-MMDYPTFLGLPAPNPKVCGCDRLSEAOQRTTGLRALAGLVNAGSGSEHCY 346  
OY 121 VIVAL-----GPDGSGAGYELMATDVTDPASCLEKFNENYAAGLPVRDY 165  
Db 347 DLYRLHSCADPTGCGTGPDPARAMDY---QACTE-INLTFASNNTDMF 391

RESULT 11  
US-10-233-745-6  
; Sequence 6, Application US/10233745  
; Publication No. US20030027282A1  
; GENERAL INFORMATION:  
; APPLICANT: Huber, Brigitte T.  
; APPLICANT: Underwood, Robert  
; TITLE OF INVENTION: QUIESCENT CELL DIPEPTIDYL PEPTIDASE: A NOVEL CYTOPLASMIC SERI  
; FILE REFERENCE: 50420 CIP  
; CURRENT APPLICATION NUMBER: US/10/233,745  
; CURRENT FILING DATE: 2002-09-03  
; PRIOR APPLICATION NUMBER: US/09/388,413  
; PRIOR FILING DATE: 1999-09-01  
; PRIOR APPLICATION NUMBER: 08/944,265  
; PRIOR FILING DATE: 1997-10-06  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 537  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-233-745-6

Query Match  
Best Local Similarity 25.9%; Score 79.5; DB 15; Length 537;  
Best Local Similarity 25.9%; Pred. No. 3.8;  
Matches 44; Conservative 27; Mismatches 58; Indels 41; Gaps 12;

OY 15 ODAWKHLQKL-VEENTDLIKATYKNDPVWGNDP-TGCVGTAAQNLNDEKKNVEMFPMNN 72  
Db 221 REAFROIKDLFLOGAIDTVR-----W-EFGTC-----OPLS-DEKDLTQLFPMFARN 264  
OY 73 ADVYQHTFEKATPDKMYGKKNENAITTYOTE---DGOVLTDLVAFS-----DDNCCY 120  
Db 265 AFTVLA-MMDYPTFLGLPAPNPKVCGCDRLSEAOQRTTGLRALAGLVNAGSGSEHCY 323  
OY 121 VIVAL-----GPDGSGAGYELMATDVTDPASCLEKFNENYAAGLPVRDY 165  
Db 324 DLYRLHSCADPTGCGTGPDPARAMDY---QACTE-INLTFASNNTDMF 368

RESULT 12  
US-10-103-196-17  
; Sequence 17, Application US/10103196  
; Publication No. US20030050466A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: TM4SF Polynucleotides, Polypeptides, and Antibodies  
; FILE REFERENCE: PT008P1  
; CURRENT APPLICATION NUMBER: US/10/103,196  
; CURRENT FILING DATE: 2002-03-22







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2003, 12:15:58 ; Search time 17.0362 Seconds  
(without alignments)  
1199.017 Million cell updates/sec

Title: US-10-087-195-6  
Perfect score: 952  
Sequence: 1 DKPYMADEAANGEHQDAWKH.....NEYAAGLPVRDYVTSDCIPE 172

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues  
Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*  
18: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	952	100.0	172	14	US-10-085-572-6 Sequence 6, Appl1
2	952	100.0	172	14	US-10-087-195-6 Sequence 6, Appl1
3	630	66.2	171	14	US-10-085-572-7 Sequence 7, Appl1
4	630	66.2	171	14	US-10-087-195-7 Sequence 7, Appl1
5	267.5	28.1	182	14	US-10-085-572-8 Sequence 8, Appl1
6	267.5	28.1	182	14	US-10-087-195-8 Sequence 8, Appl1
7	79.5	8.4	490	15	US-10-012-542-250 Sequence 250, App
8	79.5	8.4	490	15	US-10-012-542-518 Sequence 518, App
9	79.5	8.4	492	14	US-10-068-134-1 Sequence 1, Appl1
10	79.5	8.4	537	15	US-10-233-745-6 Sequence 6, Appl1
11	79.5	8.4	537	15	US-10-233-745-6 Sequence 6, Appl1
12	77.5	8.1	268	15	US-10-103-196-17 Sequence 17, Appl1
13	77.5	8.1	969	11	US-09-961-403-6 Sequence 13509, A
14	77.5	8.1	1806	15	US-10-156-761-13509 Sequence 14, Appl1
15	77	8.1	481	10	US-09-989-545-14

16	76	8.0	221	9	US-09-728-914-4	Sequence 4, Appl1
17	76	8.0	1011	15	US-10-128-714-8129	Sequence 8129, Ap
18	75.5	7.9	468	9	US-09-726-643-51	Sequence 51, Appl
19	75.5	7.9	468	10	US-09-978-295A-337	Sequence 337, App
20	75.5	7.9	468	10	US-09-978-697-337	Sequence 337, App
21	75.5	7.9	468	10	US-09-978-192A-337	Sequence 337, App
22	75.5	7.9	468	10	US-09-999-832A-337	Sequence 337, App
23	75.5	7.9	468	11	US-09-978-189-337	Sequence 337, App
24	75.5	7.9	468	11	US-09-978-608A-337	Sequence 337, App
25	75.5	7.9	468	11	US-09-978-585A-337	Sequence 337, App
26	75.5	7.9	468	11	US-09-978-191A-337	Sequence 337, App
27	75.5	7.9	468	11	US-09-978-403A-337	Sequence 337, App
28	75.5	7.9	468	11	US-09-978-564A-337	Sequence 337, App
29	75.5	7.9	468	11	US-09-999-833A-337	Sequence 337, App
30	75.5	7.9	468	11	US-09-981-815A-337	Sequence 337, App
31	75.5	7.9	468	11	US-09-978-824-337	Sequence 337, App
32	75.5	7.9	468	11	US-09-918-985A-337	Sequence 337, App
33	75.5	7.9	468	11	US-09-978-423A-337	Sequence 337, App
34	75.5	7.9	468	11	US-09-978-193A-337	Sequence 337, App
35	75.5	7.9	468	11	US-09-999-830A-337	Sequence 337, App
36	75.5	7.9	468	11	US-09-978-757A-337	Sequence 337, App
37	75.5	7.9	468	11	US-09-978-187B-337	Sequence 337, App
38	75.5	7.9	468	11	US-09-978-643A-337	Sequence 337, App
39	75.5	7.9	468	12	US-09-978-375A-337	Sequence 337, App
40	75.5	7.9	468	14	US-10-042-141-51	Sequence 51, Appl
41	75.5	7.9	468	15	US-10-017-081A-337	Sequence 337, App
42	75.5	7.9	468	15	US-10-167-749-337	Sequence 337, App
43	75.5	7.9	468	15	US-10-013-921A-337	Sequence 337, App
44	75.5	7.9	468	15	US-10-013-929A-337	Sequence 337, App
45	75.5	7.9	468	15	US-10-016-177A-337	Sequence 337, App

ALIGNMENTS

RESULT 1  
US-10-085-572-6  
: Sequence 6, Application US/10085572  
: Publication No. US20020151499A1  
: GENERAL INFORMATION:  
: APPLICANT: Nuttall, Patricia, Ann  
: TITLE OF INVENTION: Treatment of Conjunctivitis  
: FILE REFERENCE: 2488-1-003  
: CURRENT APPLICATION NUMBER: US/10/085, 572  
: PRIOR FILING DATE: 2002-02-27  
: PRIOR APPLICATION NUMBER: PCT/SB00/03282  
: PRIOR FILING DATE: 2000-08-24  
: PRIOR APPLICATION NUMBER: 9920674..0  
: NUMBER OF SEQ ID NOS: 8  
: SOFTWARE: FastSeq for Windows Version 4.0  
: SEQ ID NO: 6  
: LENGTH: 172  
: TYPE: PRT  
: ORGANISM: FS-HBP1  
US-10-085-572-6

Query Match 100.0%; Score 952; DB 14; Length 172;  
Best Local Similarity 100.0%; Pred. No. 4.4e-97;  
Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DKPYMADEAANGEHQDAWKHKLIVERNYLLIKATYKNDPVKGNDFCVGTAAQNLEDE 60  
DB 1 DKPYMADEAANGEHQDAWKHKLIVERNYLLIKATYKNDPVKGNDFCVGTAAQNLEDE 60  
QY 61 KNEAEMFEMNNADTYQHFFEKATPDKMGYKKNENATYQTEDEGCVTLVLAFFSDNDCY 120  
DB 61 KNEAEMFEMNNADTYQHFFEKATPDKMGYKKNENATYQTEDEGCVTLVLAFFSDNDCY 120  
QY 121 VITALGPDGSGAGYEIMATDYTDVPASCKLEKFNENYAAGLPVRDYVTSDCIPE 172  
DB 121 VITALGPDGSGAGYEIMATDYTDVPASCKLEKFNENYAAGLPVRDYVTSDCIPE 172

[illegible]

RC STRAIN-168;  
RA MEDLINE-21242727; PubMed-11344136;  
RX SCHULTZ A.C., NYGAARD P., SAXILD H.H.;  
RT "Functional analysis of 14 genes that constitute the purine catabolic  
RT pathway in *Bacillus subtilis* and evidence for a novel regulon  
RT controlled by the PUCR transcription activator.";  
RL J. Bacteriol. 183:3293-3302(2001).  
CC -I- FUNCTION: Oxidizes hypoxanthine and xanthine to uric acid.  
CC -I- CATALYTIC ACTIVITY: Xanthine + NAD(+) + H2O = urate + NADH.  
CC -I- COFACTOR: Molybdopterin (By similarity).  
CC -I- PATHWAY: Purine catabolism.  
CC -I- SUBUNIT: Could be composed of four subunits: pucA, pucC, pucD and  
CC pucE.  
CC -I- INDUCTION: Expression is very low in excess nitrogen (glutamate  
CC plus ammonium) and is induced during limiting-nitrogen conditions  
CC (glutamate). Expression decreases when allantoin is added during  
CC limiting-nitrogen conditions.  
CC -I- SIMILARITY: BELONGS TO THE XANTHINE DEHYDROGENASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: 299120; CAB15238.1; -  
 DR PIR: B70017; B70017.  
 DR SUDLILIST: B613989; PUCD.  
 DR InterPro: IPR000674; Aldxan\_dh\_C.  
 DR Pfam: PF02738; Ald\_Xan\_dh\_C2; 1.  
 DR Pfam: PF01315; Ald\_Xan\_dh\_C; 1.  
 KM Oxidoreductase: NAD, Molybdenum: Purine metabolism; Complete proteome.  
 SQ SEQUENCE 745 AA; 80433 MW; CA3EC5F0094EFD CRC64;

Query Match 9.0%; Score 85.5; DB 1; Length 745;  
 Best Local Similarity 26.2%; Pred. No. 5.9;  
 Matches 32; Conservative 22; Mismatches 41; Indels 27; Gaps 6;

QY 22 QKVEENYDLIKATYKNDPVGNDFTCVTAQNLNDEKNEVMEFMNN-----A 73  
 DB 106 EEIAEAALEIIOVEYKELEWDSPEKALRPNAQRLHEDGILHRAF-FSNGDVEEGFQAS 164  
 QY 74 DTVYQHTFEKATPDKMGYNKKNATVYQTEGQVLTVDVLAFSODNCVYVYALGPDGSGAG 133  
 DB 165 DTVFEETYE--LPRQMTY-----METEGVAVPE---DDGFTMTA---GTQHG 206  
 QY 134 YE 135  
 DB 207 YK 208

RESULT 6  
 ID AMY\_STRLI STANDARD; PRT; 919 AA.  
 AC Q05884;  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 01-FEB-1995 (Rel. 31, Last annotation update)  
 DE Alpha-amylose precursor (EC 3.2.1.1) (1,4-alpha-D-glucan  
 DE glucanohydrolase).  
 GN AMY.  
 OS Streptomyces lividans.  
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1916;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 34-43.  
 RC STRAIN=FK24;  
 RX MEDLINE=93144336; PubMed=8424949;  
 RA Tsao L.-S., Lin L.-L., Chen J.-C., Chen J.-H., Hsu W.-H.;  
 RT "Cloning and characterization of an alpha-amylose gene from  
 RT Streptomyces lividans.";  
 RL Biochim. Biophys. Acta 1171:255-262(1993).  
 RN [2]  
 RP ERRATUM.  
 RX MEDLINE=93250042; PubMed=8485150;  
 RA Tsao L.-S., Lin L.-L., Chen J.-C., Chen J.-H., Hsu W.-H.;  
 RL Biochim. Biophys. Acta 1173:119-119(1993).  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
 CC linkages in oligosaccharides and polysaccharides.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: X70255; CAA49759.1; -

DR HSSP: P10529; 7TAA.  
 DR InterPro: IPR006589; Alp\_amy1\_cat\_sub.  
 DR InterPro: IPR006048; Alpha\_amy1\_C.  
 DR InterPro: IPR006047; Alpha\_amy1\_cat.  
 DR InterPro: IPR005323; PUD.  
 DR Pfam: PF00128; alpha-amylose; 1.  
 DR Pfam: PF02806; alpha-amylose\_C; 1.  
 DR Pfam: PF03714; PUD; 2.  
 DR SMART: SM00642; Amy; 1.  
 KM Hydrolyase; Glycosidase; Carbohydrate metabolism; Signal.  
 FT SIGNAL 1 33  
 FT CHAIN 34 919 ALPHA-AMYLASE.  
 FT ACT\_SITE 312 312 BY SIMILARITY.  
 FT ACT\_SITE 316 316 BY SIMILARITY.  
 FT ACT\_SITE 417 417 BY SIMILARITY.  
 SQ SEQUENCE 919 AA; 100642 MW; 035B59FB80F0BD10 CRC64;

Query Match 8.6%; Score 81.5; DB 1; Length 919;  
 Best Local Similarity 27.0%; Pred. No. 17;  
 Matches 34; Conservative 11; Mismatches 32; Indels 49; Gaps 7;

QY 23 KIVEENYDLIKA-----TKNDPV-----WGNDPCV-----GTAQ 54  
 DB 101 KGLTEKLDYIKGLGTTSTIMAPIFKNQPVGGTKDASAGYHGWTIDTVDPHFGT--- 157  
 QY 55 NLNDEKN-----VEAMFMENNADTVYQTEKAT-PDKMGYNKKNATVYQTE 104  
 DB 158 --NKDKNLISKAKAKMKYF-----DVTNHTADVYVEKSYLSKGFPIYTKD 209  
 QY 105 GQVLTG 110  
 DB 210 GQPFDD 215

RESULT 7  
 ID HTRE\_ECOLI STANDARD; PRT; 865 AA.  
 AC P33129;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Outer membrane usher protein htre precursor (Heat shock protein E).  
 GN HTRE OR B0139.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE=93352405; PubMed=8102362;  
 RA Raina S., Missiakas D., Baird L., Kumar S., Georgopoulos C.;  
 RT "Identification and transcriptional analysis of the Escherichia coli  
 RT htre operon which is homologous to pap and related pilin operons.";  
 RL J. Bacteriol. 175:5009-5021(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / W3110;  
 RX MEDLINE=94261430; PubMed=8202364;  
 RA Fujita N., Mori H., Yura T., Ishihama A.;  
 RT "Systematic sequencing of the Escherichia coli genome: analysis of  
 RT the 2.4-4.1 mb (110,917-193,643 bp) region.";  
 RL Nucleic Acids Res. 22:1637-1639(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1234-1244(1997).

[4]  
RN SEQUENCE OF 30-33.  
RP Raina S.:  
RL Submitted (DEC-1994) to the SWISS-PROT data bank.  
CC -1- FUNCTION: PROBABLE PORIN-LIKE PROTEIN NECESSARY FOR THE ASSEMBLY  
CC OF A PILIN-TYPE PROTEIN.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane  
CC (By similarity).  
CC -1- INDUCTION: By heat shock.  
CC -1- SIMILARITY: BELONGS TO THE FIMBRIAL EXPORT USHER FAMILY.  
CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A  
CC FRAMESHIFT IN POSITION 481 WHICH CREATES TWO SEPARATE ORFS.  
CC -1- CAUTION: REF.1 SEQUENCE WAS INCORRECT IN POSITION 861 ONWARD DUE  
CC TO CLONING ARTIFACT.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: L00680; AAA23721.1; ALT-SEQ.  
DR EMBL: D26562; NOT\_ANNOTATED\_CDS.  
DR EMBL: AE000123; AAC73250.1; -.  
DR PIR: C64737; C64737.  
DR Ecocore: EG11972; htre.  
DR InterPro: IPR000015; Fimb\_usher.  
DR Pfam: PF00577; Usher; 1.  
DR Prosite: PS01151; FIMBRIAL-USHER; 1.  
KW Outer membrane; Transmembrane; Fimbria; Transport; signal; Heat shock;  
KW Complete proteome.  
FT SIGNAL 1 29  
FT CHAIN 30 865 OUTER MEMBRANE USHER PROTEIN HTRE.  
FT DISULFID 838 862  
FT CONFLICT 31 51 S -> T (IN REF. 1).  
FT CONFLICT 810 811 OG -> HR (IN REF. 1).  
FT CONFLICT 849 849 E -> P (IN REF. 1).  
SQ SEQUENCE .865 AA; 95499 MW; F1748B551E45AAE CRC64;  
  
Query Match 8.58; Score 80.5; DB 1; Length 865;  
Best Local Similarity 19.38; Pred. No. 19;  
Matches 46; Conservative 33; Mismatches 60; Indels 99; Gaps 12;  
  
QY 5 WADEAENGEDDAWKHLQKLVENYDLIKATKYPWNGDFTCVGTAQNLNDEKNVE 64  
Db 530 WSDVWASGQNR-----SNYSI---GYSNSTSGSISV---SAQRSMNEDGDTDD 572  
QY 65 AMFW-----FMNADTVYQ-----78  
Db 573 SVYLSFTPIETKLGTEORTSGFOSIDTQISDFKNNQLNVSSSGSDNARVSYNTG 632  
QY 79 HTEFKATPDKNY--GY-----AKENA--ITYOTEDGOVL--TDVLASDD 117  
Db 633 YTNMKAKSDLSYGVGYSYSPMGCTLAGSISANDNSRQVSLSTDGFFVLHSGGLTFSND 692  
QY 118 N-----CYVYALGPGSGAGY-----ELWATDYTPVASCLEKFNFEYAAGLPVRY 164  
Db 693 SFSQSDTLAVYQAFAGAGARINYGNSITIDRWGIVT-----SALSPYHENRIALDINDL 746  
  
RESULT 8  
ID ENVF\_SALTY STANDARD; PRT; 262 AA.  
AC 056032;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DE 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Probable lipoprotein envf precursor.  
GN ENVF OR STM1240.  
OS Salmonella typhimurium.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Salmonella.  
OX NCBI\_Taxid:602;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 14028;  
RX MEDLINE-95394809; Pubmed-7665482;  
RA Gunn J.S., Alpuche-Aranda C.M., Loomis W.P., Belden W.J., Miller S.I.;  
RT "Characterization of the salmonella typhimurium pagC/pagD chromosomal  
RT region";  
RL J. Bacteriol. 177:5040-5047(1995).  
RN [2]  
RN SEQUENCE FROM N.A.  
RC STRAIN-LT2 / SSGC1412 / ATCC 700720;  
RX MEDLINE-21534948; Pubmed-11677609;  
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
RA Courtney L., Potwolk S., Ali J., Dante M., Du P., Hou S., Layman D.,  
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,  
RA Waterston R., Wilson R.K.;  
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium  
RT LT2";  
RL Nature 413:852-856(2001).  
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor  
CC (Probable).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U31849; AAA82997.1; -.  
DR EMBL: AE008754; AAL20169.1; -.  
DR StyGene; SG10627; envf.  
DR Prosite; PS00013; PROKAR\_LIPOPROTEIN; 1.  
KW Membrane; Lipoprotein; Signal; Complete proteome.  
FT SIGNAL 1 25  
FT CHAIN 26 262 PROBABLE LIPOPROTEIN ENVF.  
FT LIPID 26 26 N-ACYL DIGLYCERIDE (POTENTIAL).  
FT CONFLICT 223 223 D -> A (IN REF. 1).  
FT CONFLICT 245 262 QRRKAAQNMNIQDTFRK -> PA (IN REF. 1).  
SQ SEQUENCE 262 AA; 29572 MW; 2BC6042085839143 CRC64;  
  
Query Match 8.48; Score 80; DB 1; Length 262;  
Best Local Similarity 23.78; Pred. No. 5.1;  
Matches 27; Conservative 21; Mismatches 46; Indels 20; Gaps 5;  
  
QY 38 NDPVWGN---DFTCVGTAQNLNDEKNVEAMFPMNADTV-----YQHTFEKAT 85  
Db 110 NNGVWGQSPVNLITFTIME--QKQELKALKF---DSINFSIIPERIGETIRAN 164  
QY 86 PDKMYGNKENVATVOTE---DGOVLVDYLAFSQDNCYVYALCPDSSGAGYEL 136  
Db 165 ASGIISVTESDIVYRAEIAHNGEFYDITTAKTARAVATLTKDGSIAGYEI 218  
  
RESULT 9  
ID ARJ2\_ANASP STANDARD; PRT; 387 AA.  
AC Q8YF9;  
DT 15-SEP-2003 (Rel. 42, Created)  
DT 15-SEP-2003 (Rel. 42, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Arginine biosynthesis bifunctional protein argJ 2 [Includes: Glutamate  
DE N-acetyltransferase (EC 2.3.1.35) (Ornithine acetyltransferase)  
DE (Ornithine transacetylase) (OATase); Amino-acid acetyltransferase  
DE (EC 2.3.1.1) (N-acetylglutamate synthase) (AGS)] [Contains: Arginine  
DE biosynthesis bifunctional protein argJ2 alpha chain; Arginine  
DE biosynthesis bifunctional protein argJ2 beta chain].  
GN ARGJ2 OR ALR4235.  
OS Anabaena sp. (strain PCC 7120).

CC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 OX NCBI\_TaxID=103690;  
 RN [1]  
 RP MEDLINE=21595285; PubMed=11759840;  
 RA Kaneko T., Nakamura Y., Molk G.P., Kuritz T., Sasamoto S.,  
 RA Watanabe A., Iriyuhui M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Shimp S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;  
 RT "Complete genomic sequence of the filamentous nitrogen-fixing  
 cyanobacterium Anabaena sp. strain PCC 7120.";  
 RL DNA Res. 8:205-213(2001).  
 CC -1- FUNCTION: Catalyzes two activities which are involved in the  
 CC cyclic version of arginine biosynthesis: the synthesis of  
 CC acetylglutamate from glutamate and acetyl-CoA, and of ornithine by  
 CC transacylation between acetylornithine and glutamate (By  
 CC similarity).  
 CC -1- CATALYTIC ACTIVITY: N(2)-acetyl-L-ornithine + L-glutamate = L-  
 CC ornithine + N-acetyl-L-glutamate.  
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + L-glutamate = CoA + N-acetyl-L-  
 CC glutamate.  
 CC -1- PATHWAY: Arginine biosynthesis; first step.  
 CC -1- SUBUNIT: Heterotrimer of two alpha and two beta chains (By  
 CC similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC -1- MISCELLANEOUS: Some bacteria possess a monofunctional argJ, i.e.,  
 CC capable of catalyzing only the fifth step of the arginine  
 CC biosynthetic pathway.  
 CC -1- SIMILARITY: Belongs to the argJ family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AP003595; BAB75934.1; -  
 DR PIR: AD2335; AD2335.  
 DR HAMAP: MF\_01106; -; 1.  
 DR InterPro: IPR002813; ArgJ.  
 DR Pfam: PF01960; ArgJ; 1.  
 DR ProDom: PD004193; ArgJ; 1.  
 DR TIGRPFAMs: TIGR00120; ArgJ; 1.  
 DR TrEMBL: TIGR00120; ArgJ; 1.  
 KM Acyltransferase; Multifunctional enzyme; Transferase;  
 FT CHAIN 1 179  
 FT ARGININE BIOSYNTHESIS BIFUNCTIONAL  
 FT PROTEIN ARGJ2 ALPHA CHAIN (BY  
 FT SIMILARITY).  
 FT ARGININE BIOSYNTHESIS BIFUNCTIONAL  
 FT PROTEIN ARGJ2 BETA CHAIN (BY  
 FT SIMILARITY).  
 FT CHAIN 180 387  
 FT ARGININE BIOSYNTHESIS BIFUNCTIONAL  
 FT PROTEIN ARGJ2 ALPHA CHAIN (BY  
 FT SIMILARITY).  
 FT SITE 179 180 CLEAVAGE (NONHYDROLYTIC) (BY SIMILARITY).  
 FT SEQUENCE 387 AA; 41145 MW; B4A340DC3FF3013D CRC64;  
 SO  
 Query Match 8.4%; Score 80; DB 1; Length 387;  
 Best Local Similarity 25.5%; Pred. No. 8.2;  
 Matches 36; Conservative 23; Mismatches 50; Indels 32; Gaps 8;

RESULT 10  
 SYSCT\_YEAST  
 ID SYSCT\_YEAST STANDARD; PRT; 462 AA.  
 AC P07284;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Seryl-tRNA synthetase, cytoplasmic (EC 6.1.1.11) (Serine--tRNA  
 DE ligase) (SERS).  
 GN SSI1 OR SRS OR YDR023W OR YD9813.01.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP MEDLINE=87174725; PubMed=3031581;  
 RA Weygand-Burzevic I., Johnson-Burke D., Soell D.;  
 RT "Cloning and characterization of the gene coding for cytoplasmic  
 RT seryl-tRNA synthetase from Saccharomyces cerevisiae.";  
 RL Nucleic Acids Res. 15:1887-1904(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RA Bowman S., Barrell B.G., Rajandream M.A.;  
 RN Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.  
 RP [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97051598; PubMed=8896275;  
 RA Elde L.G., Sander C., Prydz H.;  
 RT "Sequencing and analysis of a 35.4 kb region on the right arm of  
 RT chromosome IV from Saccharomyces cerevisiae reveal 23 open reading  
 RT frames.";  
 RL Yeast 12:1085-1090(1996).  
 RN [4]  
 RP SEQUENCE OF 451-462 FROM N.A.  
 RC STRAIN=S288C;  
 RA Polley L.S., Fox T.D.;  
 RL Submitted (XXX-1993) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ATP + L-serine + tRNA(Ser) = AMP + diphosphate  
 CC + L-seryl-tRNA(Ser).  
 CC -1- SUBUNIT: Homodimer.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X04884; CA28572.1; -  
 DR EMBL: Z47814; CAAB7802.1; -  
 DR EMBL: X95966; CA65216.1; -  
 DR EMBL: Z74319; CA938844.1; -  
 DR EMBL: L15408; AAC37412.1; -  
 DR PIR: S50930; YSBYC.  
 DR HSSP: P34945; ISER.  
 DR SGD: S0002430; SSSI.  
 DR GO: GO:0005737; C:cytoplasm; IDA.  
 DR GO: GO:0004828; F:serine-tRNA ligase activity; IDA.  
 DR InterPro: IPR002314; tRNA-synt\_2b.  
 DR InterPro: IPR002317; tRNA-synt\_ser.  
 DR InterPro: IPR006195; tRNA\_ligase\_II.  
 DR Pfam: PF02403; Seryl\_tRNA\_N; 1.  
 DR Pfam: PF00587; tRNA-synt\_2b; 1.  
 DR PRINTS: PR00981; TRNASYNTHSR.  
 DR TIGRPFAMs: TIGR00414; serS; 1.  
 DR PROSITE: PS50862; AA\_TRNA\_LIGASE\_II; 1.

FW	AMINOACYL-TRNA synthetase: Protein biosynthesis. Ligase; ATP-binding.
FT	CONFLICT 224 224 L -> P (IN REF. 1).
SO	SEQUENCE 462 AA; 53309 MW; BBED6EA47B6F547B CRC64;
Qy	Query Match 8.4%; Score 80; DB 1; Length 462;
Db	Best Local Similarity 24.6%; Pred. NO. 10;
Matches	33; Conservative 17; Mismatches 40; Indels 44; Gaps 7;
Qy	13 EHQAMKHLQLV---EENYDLIATYKNDYWGNDYFGYTAONLE--DEKNVEM 66
Db	308 EPEKSWEEFEKEMISYSEEFYSLKPYR-----YGVSGELNNAAKKYDLEAM 357
Qy	67 FFMF-----NNADIVYQTFE-----KATPKMYG-----KKNAT----- 99
Db	358 FPYGKEKVELYSCSNCIDYGRNLEIRGCIKKMGREKKYVHCLNSTLAATGRALCILE 417
Qy	100 -YQEDGOVLTDLV 112
Db	418 NYQTEDEGLVPEVL 431
RESULT 11	
DP22_HUMAN	
ID	DP22_HUMAN STANDARD: PRT: 492 AA.
AC	Q9UHL4; Q969Y4;
DT	16-OCT-2001 (Rel. 40, Created)
DT	28-FEB-2003 (Rel. 41, Last sequence update)
DT	15-SEP-2003 (Rel. 42, Last annotation update)
DE	Dipeptidyl-peptidase II precursor (EC 3.4.14.2) (DPP II) (Dipeptidyl aminopeptidase II) (Quiescent cell proline dipeptidase) (Dipeptidyl peptidase 7).
GN	DPP2 OR DPP7 OR QPP.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX	MEDLINE=20036542; PubMed=10567372;
RA	Underwood R., Chiravuri M., Lee H., Schmitz T., Kabacnel A.K.,
RA	Yardley Y., Huber B.T.;
RT	"Sequence, purification, and cloning of an intracellular serine
RL	protease, quiescent cell proline dipeptidase.";
RN	J. Biol. Chem. 274:34053-34058(1999).
RP	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Brain;
RX	MEDLINE=22388257; PubMed=12477932;
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA	Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA	Stapleton M., Soares M.B., Bonaldo A.F., Casavant T.L., Schetz T.E.,
RA	Brownstein M.J., Udell T.B., Toshlyuk S., Carlini P., Prange C.,
RA	Rata S.S., Loguelfano N.A., Peters G.J., Abraham R.D., Mullany S.J.,
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,
RA	Villalona D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Fahney J., Heltan E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grilwood J., Schmitz J., Myers R.M.,
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA	Schmerch A., Schein J.E., Jones S.J.W., Marra M.A.;
RT	"Generation and initial analysis of more than 15,000 full-length
RL	human and mouse cDNA sequences.";
RT	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RP	[3]
RP	CHARACTERIZATION
RX	MEDLINE=20521747; PubMed=11067927;
RA	Chiravuri M., Aggarberes F., Mathieu S.L., Lee H., Huber B.T.;
RT	"Vestibular localization and characterization of a novel

```

RT post-proline-cleaving aminodipeptidase, quiescent cell proline
RL dipeptidase.*.
CC J. Immunol. 165:5695-5702(2000).
CC
CC -1- FUNCTION: Plays an important role in the degradation of some
CC oligopeptides.
CC CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Xbb-1-
CC Xcc, preferentially when Xbb is Ala or Pro. Substrates are
CC oligopeptides, preferentially tripeptides.
CC
CC -1- SUBUNIT: Homodimer (Probable).
CC -1- SUBCELLULAR LOCATION: Lysosomal and intracellular vesicles.
CC -1- PTM: N-glycosylated.
CC -1- MISCELLANEOUS: HAS AN UNUSUALLY BROAD PH RANGE IN THAT IT IS ABLE
CC TO CLEAVE SUBSTRATE MOLECULES AT ACIDIC PH AS WELL AS AT NEUTRAL
CC PH.
CC
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S28.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on les-
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.sdb.ch).
CC -----
CC DR EMBL; AF154502; AAF12747.1; -
CC DR EMBL; BC011907; AAH11907.1; -
CC DR EMBL; BC016961; AAH16961.1; -
CC DR MEROPS; S28.0062;
CC
CC DR Genew; HGNC:14892; DPP7.
CC DR InterPro; IPR000073; A/D_hydrolase.
CC DR InterPro; IPR000379; Ser_estrs_site.
CC DR Pfam; PF00561; abhydrolase_1.
CC
CC KW Hydrolase; Serine protease; Aminopeptidase; Lysosome; Glycoprotein;
CC Signal; zymogen.
CC
CC FT SIGNAL 1 21 POTENTIAL.
CC FT PROPEP 22 25 POTENTIAL.
CC FT CHAIN 26 492 DIPEPTIDYL-PEPTIDASE II.
CC FT ACCT_SITE 162 162 CHARGE RELAY SYSTEM (POTENTIAL).
CC FT ACCT_SITE 418 418 CHARGE RELAY SYSTEM (POTENTIAL).
CC FT ACCT_SITE 443 443 CHARGE RELAY SYSTEM (POTENTIAL).
CC FT CARBOHND 50 50 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHND 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHND 315 315 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHND 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHND 363 363 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHND 428 428 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CONFLICT 71 71 I -> T (IN REF. 1).
CC
CC SO SEQUENCE 492 AA; 54327 MW; 58C4BD0B87F10C4 CMC64;

```

DE Phosphatidylinositol 3-kinase tor1 (EC 2.7.1.137) (PI3-kinase)  
 DN (Ptdins-3-kinase) (PI3K).  
 GN TOR1 OR SPC30D10.10C.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 CC Schizosaccharomyces.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sounos J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Haldal J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skellon J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Voickert G., Aert R., Robben J., Gymnopoulos B.,  
 RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,  
 RA Eber P., Zimmermann W., Wedler H., Wambutt R., Pinnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hunt S.M.,  
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Rhode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Sapatovsky G.V., Ussery D., Barrell B.G., Nurse P.;  
 RA "The genome sequence of Schizosaccharomyces pombe.";  
 RT Nature 415:871-880(2002).  
 RN [2]  
 RP IDENTIFICATION, FUNCTION, AND INDUCTION.  
 RX MEDLINE=21269225; PubMed=11096119;  
 RA Weisman R., Choder M.;  
 RT "The fission yeast TOR homolog, tor1+, is required for the response to  
 RT starvation and other stresses via a conserved serine.";  
 RT J. Biol. Chem. 276:7027-7032(2001).  
 CC -1- FUNCTION: phosphatidylinositol 3-kinase homolog required for G1  
 CC progression and entry into stationary phase. Also required for the  
 CC onset of meiosis and sporulation under nitrogen and carbon  
 CC starvation conditions.  
 CC -1- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol = ADP +  
 CC 1-phosphatidyl-1D-myo-inositol 3-phosphate.  
 CC -1- INDUCTION: By nitrogen and/or carbon starvation, cold, osmotic and  
 CC oxidative stress.  
 CC -1- SIMILARITY: BELONGS TO THE PI3/PI4-KINASES FAMILY.  
 CC -1- SIMILARITY: Contains 16 HEAT repeats.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL\_297992; CAB10805.1; -;  
 DR PIR; T40186; T40186.  
 DR HSSP; P42345; 1FAP.  
 DR GeneDB SPombe: SPC30D10.10c; -;  
 DR InterPro: IPR003151; FAT.  
 DR InterPro: IPR003152; FATC.  
 DR InterPro: IPR000357; HEAT\_repeat.  
 DR InterPro: IPR000403; PI3\_PI4\_kinase.  
 DR Pfam; PF02259; FAT; 1.

DR Pfam; PF02260; FATC; 1.  
 DR PIR; T40186; T40186.  
 DR SMART; SM00146; PI3K; 1.  
 DR PROSITE; PS50077; HEAT\_REPEAT; FALSE NEG.  
 DR PROSITE; PS00915; PI3\_4\_KINASE\_1; 1.  
 DR PROSITE; PS00916; PI3\_4\_KINASE\_2; 1.  
 DR PROSITE; PS50290; PI3\_4\_KINASE\_3; 1.  
 KW Transferase; Kinase; Cell cycle; Repeat.  
 FT REPEAT 1  
 FT REPEAT 164 201 HEAT 1.  
 FT REPEAT 331 371 HEAT 2.  
 FT REPEAT 410 449 HEAT 3.  
 FT REPEAT 474 512 HEAT 4.  
 FT REPEAT 522 560 HEAT 5.  
 FT REPEAT 562 596 HEAT 6.  
 FT REPEAT 642 679 HEAT 7.  
 FT REPEAT 684 722 HEAT 8.  
 FT REPEAT 728 766 HEAT 9.  
 FT REPEAT 788 826 HEAT 10.  
 FT REPEAT 843 880 HEAT 11.  
 FT REPEAT 904 923 HEAT 12.  
 FT REPEAT 924 961 HEAT 13.  
 FT REPEAT 964 1003 HEAT 14.  
 FT REPEAT 1005 1042 HEAT 15.  
 FT REPEAT 1132 1169 HEAT 16.  
 FT DOMAIN 1987 2335 PI3K/PI4K.  
 SQ SEQUENCE 2335 AA; 266181 MW; 5DCE1CF4B8E9A4 CRC64;  
 Query Match 8.48; Score 79.5; DB 1; Length 2335;  
 Best Local Similarity 22.68; Pred. No. 78;  
 Matches 43; Conservative 29; Mismatches 55; Indels 63; Gaps 11;  
 QY 12 GENDAKKHOK---LVEENYDLIKTYKNDPVWGNDPFCVGTAAQNLNDE-----K 61  
 DB 1402 GQYKKAAGHIERHDLIVNLSSTIGSYR--AVG-----INVSQMSLELEIIDYKR 1454  
 QY 62 NVEAMFENMANNADVTVOHFEKATPDKMGYNKE-----NATYQTEGQ----V 107  
 DB 1455 NMQ---YENNIDSL-----KTMWRKLECCQKNVDVWNTLFRALVLSPODSEPMIK 1505  
 QY 108 LTVLAESD---DNCYVYIALGPDGSGAGYE-----LWATDYDVP 145  
 DB 1506 LADICRSRSDIKLSNOCITVYMGDRPSNA--YPDLSKLINPHVYVYTYLKYLMATDKNIA 1564  
 QY 146 ASCLERFNEY 155  
 DB 1565 VSELEERTSY 1574  
 RESULT 13  
 PFTA\_SCHPO  
 AC 060052; STANDARD; PRT; 294 AA.  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE Protein farnesyltransferase alpha subunit (EC 2.5.1.-) (CAAX  
 DE farnesyltransferase alpha subunit) (RAS proteins prenyltransferase  
 DE alpha) (Frase alpha).  
 GN CWP1 OR SPAPBIA10.04C.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 CC Schizosaccharomyces.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A., FUNCTION, AND SUBUNIT.  
 RC STRAIN=972;  
 RX MEDLINE=98453135; PubMed=97818174;  
 RA Atelliano M., Coll P.M., Yang W., Duran A., Tamanoi F., Perez P.;  
 RT "Characterization of the geranylgeranyl transferase type I from  
 RT Schizosaccharomyces pombe.";  
 RT Mol. Microbiol. 29:1357-1367(1998).  
 RN [2]





RA MEDLINE-97064242; PubMed-8906861;  
RA Zhong M., Benjannet S., Lazure C., Munzer S., Seldah N.G.;  
RT "Functional analysis of human PACE4-A and PACE4-C isoforms:  
RT identification of a new PACE4-CS isoform";  
RL FEBS Lett. 396:31-36(1996).  
RN [8]  
RP CHARACTERIZATION.  
RX MEDLINE-99233559; PubMed-10215603;  
RX Susic J.F., Moehring J.M., Innocencio N.M., Luchini J.W.,  
RA Moehring J.J.;  
RT "Endoprotease PACE4 is Ca<sup>2+</sup>-dependent and temperature-sensitive and  
RT can partly rescue the phenotype of a furin-deficient cell strain.",  
RL Biochem. J. 339:639-647(1999).  
RN [9]  
RP PROCESSING.  
RX MEDLINE-98408849; PubMed-9738469;  
RX Nagahama M., Taniguchi T., Hashimoto E., Imamaki A., Mori K.,  
RA Tsuji A., Matsuda Y.;  
RT "Biosynthetic processing and quaternary interactions of proprotein  
RT convertase SPC4 (PACE4)".  
RL FEBS Lett. 434:155-159(1998).  
CC -1- FUNCTION: LIKELY TO REPRESENT AN ENDOPROTEASE ACTIVITY WITHIN THE  
CC CONSTITUTIVE SECRETORY PATHWAY, WITH UNIQUE RESTRICTED  
CC DISTRIBUTION IN BOTH NEUROENDOCRINE AND NON-NEUROENDOCRINE TISSUES  
CC AND CAPABLE OF CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF.  
CC -1- CATALYTIC ACTIVITY: Release of mature proteins from their  
CC proproteins by cleavage of Arg-Xaa-Yaa-Ark-1'-Zaa bonds, where Xaa  
CC can be any amino acid and yaa is Arg or Lys.  
CC -1- COFACTOR: PACE4A IS PROBABLY CALCIUM-DEPENDENT.  
CC -1- SUBUNIT: THE PACE4A-I PRECURSOR PROTEIN SEEMS TO EXIST IN THE  
CC RETICULUM ENDOPLASMIC AS BOTH A MONOMER AND A DIMER-SIZED COMPLEX  
CC WHEREAS MATURE PACE4A-I EXISTS ONLY AS A MONOMER, SUGGESTING THAT  
CC PROPEPTIDE CLEAVAGE AFFECTS ITS TERTIARY OR QUATERNARY STRUCTURE.  
CC -1- SUBCELLULAR LOCATION: PACE4A-I AND PACE4-AII ARE SECRETED. PACE4C  
CC AND PACE4CS ARE NOT SECRETED AND REMAIN PROBABLY IN ZYMOGEN FORM  
CC IN ENDOPLASMIC RETICULUM. PACE4E-I AND PACE4E-II ARE RETAINED  
CC INTRACELLULARLY PROBABLY THROUGH A HYDROPHOBIC CLUSTER IN THEIR C-  
CC TERMINUS. PACE4B MIGHT BE SECRETED.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event-Alternative splicing; Named Isoforms=8;  
CC Name-PACE4A-I; Synonyms-PACE4;  
CC IsoId=p29122-1; Sequence-Displayed;  
CC Name-PACE4A-II; Sequence-Displayed;  
CC IsoId=p29122-2; Sequence-VSP\_005436;  
CC Name-PACE4B; Synonyms-PACE4.1;  
CC IsoId=p29122-3; Sequence-VSP\_005428; VSP\_005429;  
CC Note-Probably enzymatically inactive;  
CC Name-PACE4C;  
CC IsoId=p29122-4; Sequence-VSP\_005432; VSP\_005433;  
CC Note-Probably enzymatically inactive;  
CC Name-PACE4CS;  
CC IsoId=p29122-5; Sequence-VSP\_005430; VSP\_005431;  
CC Note-Probably enzymatically inactive;  
CC Name-PACE4D;  
CC IsoId=p29122-6; Sequence-VSP\_005427; VSP\_005434; VSP\_005435;  
CC Note-Probably enzymatically inactive;  
CC Name-PACE4E-I;  
CC IsoId=p29122-7; Sequence-VSP\_005437;  
CC Name-PACE4E-II;  
CC IsoId=p29122-8; Sequence-VSP\_005436; VSP\_005437;  
CC -1- TISSUE SPECIFICITY: EACH PACE4 ISORFORM EXHIBITS A UNIQUE  
CC RESTRICTED DISTRIBUTION. PACE4A-I IS EXPRESSED IN HEART, BRAIN,  
CC PLACENTA, LUNG, SKELETAL MUSCLE, KIDNEY, PANCREAS, BUT AT  
CC COMPARATIVELY HIGHER LEVELS IN THE LIVER. PACE4A-II IS AT LEAST  
CC EXPRESSED IN PLACENTA. PACE4B WAS ONLY FOUND IN THE EMBRYONIC  
CC KIDNEY CELL LINE FROM WHICH IT WAS ISOLATED. PACE4C AND PACE4D ARE  
CC EXPRESSED IN PLACENTA. PACE4E-I IS EXPRESSED IN CEREBELLUM,  
CC PLACENTA AND PITUITARY. PACE4E-II IS AT LEAST PRESENT IN  
CC CEREBELLUM.  
CC -1- DOMAIN: THE PROPEPTIDE DOMAIN ACTS AS AN INTRAMOLECULAR CHAPERONE  
CC ASSISTING THE FOLDING OF THE ZYMOGEN WITHIN THE ENDOPLASMIC  
CC RETICULUM. ISORFORM PACE4D LACKS THE PROPEPTIDE DOMAIN.  
CC -1- SIMILARITY: Belongs to peptidase family 58.

CC -1- SIMILARITY: Contains 1 homo B/P domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M80482; AAA5998.1; -  
DR EMBL; AB001914; BAA21620.1; -  
DR EMBL; AB001898; BAA21620.1; JOINED.  
DR EMBL; AB001901; BAA21620.1; JOINED.  
DR EMBL; AB001902; BAA21620.1; JOINED.  
DR EMBL; AB001903; BAA21620.1; JOINED.  
DR EMBL; AB001904; BAA21620.1; JOINED.  
DR EMBL; AB001905; BAA21620.1; JOINED.  
DR EMBL; AB001914; BAA21621.1; -  
DR EMBL; AB001898; BAA21621.1; JOINED.  
DR EMBL; AB001900; BAA21621.1; JOINED.  
DR EMBL; AB001901; BAA21621.1; JOINED.  
DR EMBL; AB001902; BAA21621.1; JOINED.  
DR EMBL; AB001903; BAA21621.1; JOINED.  
DR EMBL; AB001904; BAA21621.1; JOINED.  
DR EMBL; AB001905; BAA21621.1; JOINED.  
DR EMBL; AB001906; BAA21621.1; JOINED.  
DR EMBL; AB001907; BAA21621.1; JOINED.  
DR EMBL; AB001908; BAA21621.1; JOINED.  
DR EMBL; AB001909; BAA21621.1; JOINED.  
DR EMBL; AB001914; BAA21623.1; -  
DR EMBL; AB001898; BAA21623.1; JOINED.  
DR EMBL; AB001900; BAA21623.1; JOINED.  
DR EMBL; AB001901; BAA21623.1; JOINED.  
DR EMBL; AB001902; BAA21623.1; JOINED.  
DR EMBL; AB001903; BAA21623.1; JOINED.  
DR EMBL; AB001904; BAA21623.1; JOINED.  
DR EMBL; AB001905; BAA21623.1; JOINED.  
DR EMBL; AB001906; BAA21623.1; JOINED.  
DR EMBL; AB001907; BAA21623.1; JOINED.  
DR EMBL; AB001908; BAA21623.1; JOINED.  
DR EMBL; AB001909; BAA21623.1; JOINED.  
DR EMBL; AB001914; BAA21624.1; -  
DR EMBL; AB001898; BAA21624.1; JOINED.  
DR EMBL; AB001900; BAA21624.1; JOINED.  
DR EMBL; AB001901; BAA21624.1; JOINED.  
DR EMBL; AB001902; BAA21624.1; JOINED.  
DR EMBL; AB001903; BAA21624.1; JOINED.  
DR EMBL; AB001904; BAA21624.1; JOINED.  
DR EMBL; AB001905; BAA21624.1; JOINED.  
Query Match 8.18; Score 77.5; DB 1; Length 969;  
Best Local Similarity 21.38; Pred. No. 41;  
Matches 39; Conservative 28; Mismatches 81; Indels 35; Gaps 7;  
QY 13 EHDAMKHLQKLEVENTDLIKATYKNDPVGNDFTCVGRAQNLNDEKRVF-AW---FM 68  
DB 139 QQQQEVRRYKRVQRSD--PQALYFNDPILWSNMWYLHCGDKNSRCSENNVQAMRGRT 195  
QY 69 FMNADTVVQHTFEKATPDKMGYKN-----ENAIY-----QTEGGVLTNDVLA 113  
DB 196 GRNVVVTILDGQIERHNPDLAPNTDSTASYDVGNIDYSPRIDSASNEKNKHTRCAGEVA 255



```

FT CARBOHYD 1732 1732 N-LINKED (GLCNAC...) (POTENTIAL).
FT VARSP LIC 878 913 XXXXXXXXXXXXXXXXXXXXXXXXXXXX ->
FT ATERSWAEGECMLVKRNNLCORRVLOOLCCKTCTFOG
FT (in isoform PCSA).
FT /Frid-VSP 005440.
FT VARSP LIC 916 1877 Missing (in isoform PCSA).
FT /Frid-VSP 005441.
SO SEQUENCE 1877 AA; 207888 MW; 890935DC60534444 CRC64;

Query Match 8.0%; Score 76.5; DB 1; Length 1877;
Best Local Similarity 25.7%; Pred. No. 1.1e+02;
Matches 29; Conservative 11; Mismatches 32; Indels 41; Gaps 6;

QY 21 LQKIVENYDLKA--TYKNDPVW-----GNDFTCVGTAAQNLNDEKRYE-AW-- 66
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
110 VKRTRKRDYDLRSRAQSTYFNDPKWPSMWMHCSDNTHPC-----QSDMNIEGAMKR 160

QY 67 -FMFMNADIVYQHTPEKATPDKMYGYNKENATITYQTEGQVLTDLVAFSDN 118
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
161 GYTGKNIIVTILDDGIERTHPDLMQNT-----DALASCDVN 196

```

Search completed: August 1, 2003, 12:15:48  
 Job time : 14.4838 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: August 1, 2003, 12:15:04 ; Search time 12.7771 Seconds  
(without alignments)  
1294.578 Million cell updates/sec

Title: US-10-087-195-6  
Perfect score: 952  
Sequence: 1 DKPYMADEANGEHQDAWKH.....NEYAAGLPVRDVYTSDCLEPE 172

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: PIR\_76:\*  
2: pir1:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88.5	9.3	919	2	hypothetical prote
2	88	9.2	648	2	hypothetical prote
3	86.5	9.2	1237	2	internalin protei
4	86.5	9.1	368	2	hypothetical prote
5	85.5	9.0	745	2	probable oxidoredu
6	84	8.8	408	2	probable cytochrom
7	83.5	8.8	133	2	hypothetical prote
8	83.5	8.8	745	2	probable transposo
9	82.5	8.7	450	2	hypothetical prote
10	82.5	8.7	475	2	beta-glucosidase (
11	82	8.6	1668	2	Iga-specific metal
12	81.5	8.6	133	2	hypothetical prote
13	81.5	8.6	512	2	hypothetical prote
14	81.5	8.6	919	2	alpha-amyrase (PC
15	81.5	8.6	1377	2	hypothetical prote
16	80.5	8.5	246	2	hypothetical prote
17	80.5	8.5	865	1	outer membrane ush
18	80	8.4	387	2	ornithine acetyl t
19	80	8.4	462	1	serine-trna ligase
20	80	8.4	601	2	casp protein - clo
21	80	8.4	932	2	PAGE4A - mouse (fr
22	79.5	8.4	2335	2	probable phosphati
23	79	8.3	580	2	alpha-glucosidase
24	79	8.3	580	2	probable alpha-glu
25	79	8.3	921	2	isooleucyl-trna syn
26	78	8.2	874	2	isooleucine trna sy
27	78	8.2	938	2	isooleucine trna sy
28	78	8.2	938	2	isooleucine trna sy
29	77.5	8.1	283	2	hypothetical prote

30	77.5	8.1	487	2	subtilisin-like pr
31	77.5	8.1	515	2	hypothetical prote
32	77.5	8.1	652	1	subtilisin-like pr
33	77.5	8.1	962	2	subtilisin-like pr
34	77.5	8.1	969	1	subtilisin-like pr
35	77.5	8.1	975	2	subtilisin-like pr
36	77.5	8.1	1157	2	fusion of alpha-gi
37	77	8.1	544	2	fructuronate reduc
38	77	8.1	754	2	lysophospholipase
39	76.5	8.0	915	2	probable propolact
40	76	8.0	298	2	conserved hypotnet
41	76	8.0	387	2	hypothetical prote
42	76	8.0	672	1	cyathadence acces
43	75.5	7.9	246	2	hypothetical prote
44	75.5	7.9	327	2	branched-chain alp
45	75.5	7.9	461	1	alpha-L-fucosidase

## ALIGNMENTS

## RESULT 1

S37786  
hypothetical protein YKL165c - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein YKL619  
C:Species: Saccharomyces cerevisiae  
C:Date: 03-May-1994 #sequence revision 03-May-1994 #text change 29-Oct-1999  
C:Accession: S37786; S37995; S44563  
R:Vandenbol, M.; Bolle, P.; Dion, C.; Portetelle, D.; Hilger, F.  
Submitted to the EMBL Data Library, September 1993  
A:Description: DNA sequencing of a 36.2 kb fragment located between the FAS1 and LAP4  
A:Reference number: S37786  
A:Accession: S37786  
A:Molecule type: DNA  
A:Residues: 1-919 <VA2>  
A:Cross-references: EMBL:Z26877; NID:q407482; PIDN:CA081489.1; PID:q407483  
A:Experimental source: strain S288C  
R:Vandenbol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.  
Submitted to the Protein Sequence Database, March 1994  
A:Reference number: S37976  
A:Accession: S37995  
A:Molecule type: DNA  
A:Residues: 1-919 <VA2>  
A:Cross-references: EMBL:Z26877; NID:q486288; PIDN:CA082007.1; PID:q486289; MIPS:YKL1  
A:Experimental source: strain S288C  
A:Title: DNA sequencing of a 36.2 kb fragment located between the FAS1 and LAP4 loci  
A:Reference number: S44563  
A:Accession: S44563  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-919 <VA3>  
A:Cross-references: EMBL:Z26877; NID:q407482; PIDN:CA081489.1; PID:q407483  
A:Experimental source: strain S288C  
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, September 199  
C:Genetics:  
A:Gene: SGD:MCD4  
A:Cross-references: SGD:S0001648; MIPS:YKL165c  
A:Map position: III

Query Match 9.3%; Score 88.5; DB 2; Length 919;

Best Local Similarity 20.1%; Pred. No. 7.2;  
Matches 34; Conservative 30; Mismatches 64; Indels 41; Gaps 7;

QY	27	ENYDIKATYKNDPVGMDFTGCVTAQNLN-----EDEKVEAMFMFMNNAD	74
DB	117	EDVSAYTKGKMGKPNFNSFQSTHTYSFGSPDILPMFKGASDPNNVDTW-----	168
QY	75	TYVOHTFEKATPKDKYGNKENATYQRED---GQVLTDLVAF---SDDNCRYVYALGP	127
DB	169	MYDHTFEDFTQST---EIDAEVFRLLDQLFNHNSLTLDYDIRODGNVFLHLGCG	223

Qy 128 DGSAGYEELMTDTD-----VPASCLEKENEYAGLPVRYTSD 168  
Db 224 DFGHSYRPSAEYDNNKYIDDIPI-LIDKVNKFFADKTAFITAD 271

## RESULT 2

474896  
hypothetical protein T4C21.260 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 05-May-2000  
C:Accession: T47896  
R:Choline, N.; Robert, C.; Brotlier, P.; Wincker, P.; Catolico, L.; Attiguenave, F.; S.  
W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.  
Submitted to the Protein Sequence Database, March 2000  
A:Reference number: 224479  
A:Accession: T47896  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-648 <CHO>  
A:Cross-references: EMBL:AL162295  
A:Experimental source: cultivar Columbia; BAC clone T4C21  
C:Genetics:  
A:Map position: 3  
A:Note: T4C21.260  
C:Superfamily: Arabidopsis thaliana hypothetical protein T4C21.260

Query Match 9.2%; Score 88; DB 2; Length 648;  
Best Local Similarity 25.3%; Pred. No. 5.1;  
Matches 38; Conservative 19; Mismatches 69; Indels 24; Gaps 6;  
Qy 2 KPWADPAANGEHODANKHLQKLEENYDLIKATYKNDPYWGNDFTCVTAONLNDEK 61  
Db 146 RAWKRDNEETAYLTAKRKID-----KLSRLDPATMGNEFLCF---KNSQOFVS 193  
Qy 62 NVEAM--FMFMNNAADYVOHTPEKATPDKMGYKKNENAIYQFEDGOVLNDFLAFSDNC 119  
Db 194 HINOMODIVMGFADGLKHLGOKETIDRKQWTVGAKLYGIPESITRVCAV-----C 248  
Qy 120 VYVYALGPDGSGAGYELMATDYT--DVPA 146  
Db 249 PVCNA--DSGSASRNKRREFYESLDVPA 276

## RESULT 3

AC1583  
internalin protein (LPXNG motif) homolog lin1204 [imported] - Listeria innocua (strain C  
C:Species: Listeria innocua  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AC1583  
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurpkat, G.; Madueno, E.; Maitounnam, A.; Ma  
ok, C.; Schluster, T.; Stimes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Weiland,  
A.; Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AC1583  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1237 <GIA>  
A:Cross-references: GB:AL552022; PIDN:GAC96435.1; PID:g16413678; GSPDB:GN00178  
A:Experimental source: strain Clp11262  
C:Genetics:  
A:Gene: lin1204

Query Match 9.2%; Score 88; DB 2; Length 1237;  
Best Local Similarity 24.8%; Pred. No. 12;  
Matches 41; Conservative 23; Mismatches 65; Indels 36; Gaps 6;  
Qy 19 KHLQKLEENYDLIKATYKNDPYWGNDFTCVTAONLN 57  
Db 775 KYIVTLNENLDGKATPKREVIVYIDAVGANTITVKEDESGNKLAEINSTLTGNVGEYS 834

Qy 58 EDERNEAMFENNNADTVYOHPEKATPDKMGYN-----KENAITYQTEDEGOVL--TDV 111  
Db 835 SSAKEILGYLTLEPNSNAGEFSLBEGTVIYIKSNVPKADIVQYTDGDELATET 894

Qy 112 LAESDNCYIYALGPDGSGAGYELMATDYTDVPASCLEKENEYA 156  
Db 895 LSGNVDENYVYTA---KTFGYEL-----LEPNSNAGEFSENA 930

## RESULT 4

S75234  
hypothetical protein slr1920 - Synechocystis sp. (strain PCC 6803)  
C:Species: Synechocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000  
C:Accession: S75234  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,  
O.; K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys  
s.  
A:Reference number: S74322; MUID:97061201; PMID:8905231  
A:Accession: S75234  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-368 <KAN>  
A:Cross-references: EMBL:D90903; GB:AB001339; NID:g1652127; PIDN:BA17148.1; PID:g165  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Superfamily: Synechocystis hypothetical protein slr1920

Query Match 9.1%; Score 86.5; DB 2; Length 368;  
Best Local Similarity 26.2%; Pred. No. 3.4;  
Matches 43; Conservative 24; Mismatches 50; Indels 47; Gaps 10;  
Qy 22 QKLEENYDL-----IKATYKNDPYWGNDFTCVTAONLNDEKNEAMFENNA 73  
Db 112 QKLYRESFNLOPGLVAIKDIRKKSIVSEKDY-----ADNIDK-----FLEIFN- 156  
Qy 74 DTVYQ-----HTPEKATPDKMGYKKNENAIYQFEDGOVLNDFLAFSDNCYIYALGPDG 129  
Db 157 PAIHQILLIRFIPVPDDMHQFDN---ILPPOFEKITLDVINFCDLN---YSL----- 205  
Qy 130 SGAGYELM-----ATDVTVPASCLEKENEYAGLPVRYTSD 168  
Db 206 DNIYDKQALVHALDETGVSLNVNKHQY-----PPREFETD 245

## RESULT 5

B70017  
probable oxidoreductase (EC 1.-.-.-) yurc [similarity] - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Dec-2002  
C:Accession: B70017  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber  
C.; Bron, S.; Brouillet, S.; Brunsch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;  
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabel, C.; Ferrari,  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gal  
tech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M  
koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau  
Y, M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanl  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Tepest, P.; Tognoni, A.; Tosato, V.; Uchiya  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida  
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis  
A:Reference number: A69580; MUID:98044033; PMID:9384377  
A:Accession: B70017  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA

A:Residues: 1-745 <KUN>  
A:Cross-references: GB:Z99120; GB:AL009126; NID:g2635613; PIDN:CAB15238.1; PID:g2635745  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yurC  
A:Superfamily: carbon-monoxide dehydrogenase molybdoprotein  
C:Keywords: oxidoreductase

Query Match 9.0%; Score 85.5; DB 2; Length 745;  
Best Local Similarity 26.2%; Pred. No. 10;  
Matches 32; Conservative 22; Mismatches 41; Indels 27; Gaps 6;

Oy 22 OKLVEENVYDLIKATYKNDPVGNDFTCVTAQNLNEDKNEVAMFMNN-----A 73  
Db 106 EETAEALDELIOVEKLEWMDSPKALRPNAQRLHEDGNLHNAF-FSNDVVEGROAS 164  
Oy 74 DTVYQHTFEKATPDKMGYKNEKNAITYQTEGQVLTDLAFSDNCTVYIALGPDGSGAG 133  
Db 165 DTVVEETVE--LPRQMHLY-----METEGVAVPE----DDGCFMTVA----GTQHG 206  
Oy 134 YE 135  
Db 207 YK 208

## RESULT 6

AF0334  
probable cytochrome c-type biogenesis protein YPO2743 [imported] - Yersinia pestis (stra  
C:Species: Yersinia pestis  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001  
C:Accession: AF0334

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tittball, R.W.; Holden, M.T.G.; Prentice, M.B.  
deno-Tarrega, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AF0334

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-408 <KUR>  
A:Cross-references: GB:AL590842; PIDN:CAC92981.1; PID:g15980720; GSPDB:GN00175  
C:Genetics:  
A:Gene: YPO2743

Query Match 8.8%; Score 84; DB 2; Length 408;  
Best Local Similarity 24.1%; Pred. No. 6.5;  
Matches 35; Conservative 21; Mismatches 60; Indels 42; Gaps 8;

Oy 13 EHODAMKHLKLVENYDLIKATYKN--DPVWGNDFTCVGTAQ-NLNEDKNEVAMFM 68  
Db 114 EQVAAMHNOVAQME-----LRARVANEHADRLSMEDVARLGLRTSLQDAGVNDMM 169  
Oy 69 F-----MNNADTVYQ--HTFEKATPDK--MGYKNEKNAITYQTEGQVLTDLAFSD 116  
Db 170 LGRVGMALNNTATQAFNAHYQLA-PDNNEVKLGVEVILRSNDPEDNQLATQMLR--- 225  
Oy 117 DNCVYIALGPDGSGAGYELMATDYTDVPASCLEKFEHYAG 158  
Db 226 -----AMYGODHTMLRAWSLLAFNAFEG 249

## RESULT 7

B81171  
hypothetical protein MMB0676 [imported] - Neisseria meningitidis (strain MC58 serogroup  
C:Species: Neisseria meningitidis  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
C:Accession: B81171

R:Teitelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.  
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
11, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Massignani, V.; Pizzi, M.  
Science 287, 1809-1815, 2000  
A:Authors: Grandt, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Va

A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
A:Reference number: A81000; MUID:20175755; PMID:10710307

A:Accession: B81171  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-133 <TEET>  
A:Cross-references: GB:AE002422; GB:AE002098; NID:g7225898; PIDN:AAF41094.1; PID:g722  
C:Genetics:  
A:Gene: MMB0676

Query Match 8.8%; Score 83.5; DB 2; Length 133;  
Best Local Similarity 25.5%; Pred. No. 1.7;  
Matches 42; Conservative 17; Mismatches 53; Indels 53; Gaps 10;

Oy 13 EHODAMKHL-----OKLVEENVYDLIKATYKNDPVG---NDFCVGTAQNLN-EDKNEV 64  
Db 3 OHIEKWOHLREEDKILAE-----VGLVQNDQEVHYEMKLKLNAPDEASGE 49  
Oy 65 AMFMNNADTVYQHTFEKATPDKMGYKNEKNAITYQTEGQVLTDLAFS---DQNCYV 121  
Db 50 FWFPMATLSTL-----PPNRSLG-----LNNNGRLTAVSILSVMTEDN--- 90  
Oy 122 IYALGPDGSGAGYELMATDYTDVPASCLEKFEHYAGLPVADVYT 166  
Db 91 -----PDIP-----QLMAQKITALNYSG-LNLYGVASPCRTCT 125

## RESULT 8

H85048  
probable transposon protein [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Feb-2001  
C:Accession: H85048

R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp  
Nature 402, 769-777, 1999  
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A:Reference number: A85001; MUID:20083488; PMID:10617198  
A:Accession: H85048

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-745 <STO>  
A:Cross-references: GB:NC\_001268; NID:g7267141; PIDN:CAB80809.1; GSPDB:GN00140  
C:Genetics:  
A:Gene: AT4G03860  
A:Map position: 4

Query Match 8.8%; Score 83.5; DB 2; Length 745;  
Best Local Similarity 25.6%; Pred. No. 16;  
Matches 42; Conservative 21; Mismatches 62; Indels 39; Gaps 10;

Oy 7 DEANGENQ-DAKHNLOKLVENYDLIKATY-----KNDFVWGNDFTCVGTAQ-NLNED 59  
Db 240 DEDIEVEPEPSWMEKTLNKPDEVTVEEYIREFENMD-FWGRYPCYETLADGLIED 298  
Oy 60 EKNVEAMFMNNADTVYQHTFEKATPDKMGYKNEKNAITYQTEGQVLTDLAFSDNC 119  
Db 299 -----YQHOFKCHLETLSY---PRATYKEBRIEFLS-TLQSLTKKL 337  
Oy 120 VYIYALGPDGSGAG-----YELMATDYTDVA-SCLEKFE 154  
Db 338 EELFGF-PSGNGTKPFERKELKDLMATSGNNVPLNSRKSQ 380

## RESULT 9

H82891  
hypothetical protein U0441 [imported] - Ureaplasma urealyticum  
C:Species: Ureaplasma urealyticum  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: H82891

R:Glass, J.I.; Letkovitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
submitted to Genbank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a

A:Reference number: A82870  
A:Accession: H82891  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-450 <GIA>  
A:Cross-references: GB:AE002140; GB:AF222894; NID:96899420; PIDN:AAF30853.1; GSPDB:GN001  
A:Experimental source: serovar 3; biovar 1  
C:Genetics:  
A:Gene: U0441  
A:Genetic code: SGC3

Query Match 8.7%; Score 82.5; DB 2; Length 450;  
Best Local Similarity 22.2%; Pred. No. 10;  
Matches 28; Conservative 25; Mismatches 44; Indels 29; Gaps 6;

QY 7 DEANGEHODAMKHLQKVEENYDLIKATYKNDPVWGNDPTCYG---TAQONLNED--- 59  
DB 19 DENSENSHIKIMKELMDLNDVKIINHSYGSP---NDPLNIGFYNFINGKEINDDKDF 75  
QY 60 EKAVEAMFPMNAD---TVQHTPEKATPDKM-----YGYKKNATYYQT 102  
DB 76 KRELKXIDTLNKDIDVNTINKKFKNYIDFISYFDIINIFSAGNSYND--IKYRN 133  
QY 103 EDGOVL 108  
DB 134 KNGHIL 139

RESULT 10  
A86811  
beta-glucosidase (EC 3.2.1.21) [imported] - Lactococcus lactis subsp. lactis (strain IL1  
C:Species: Lactococcus lactis subsp. lactis  
C>Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
C:Accession: A86811  
R:Bioclin, A.; Wancher, P.; Manger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrli  
G:Genome Res. 11, 731-753, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s  
A:Reference number: A86825; MUID:2125186; PMID:11337471  
A:Accession: A86811  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-475 <STO>  
A:Cross-references: GB:AE005176; PID:912724484; PIDN:AAK05587.1; GSPDB:GN00146  
A:Experimental source: strain IL1403  
C:Genetics:  
A:Gene: YPCA  
C:Superfamily: Agrobacterium beta-glucosidase  
C:Keywords: glycosidase; hydrolase

Query Match 8.7%; Score 82.5; DB 2; Length 475;  
Best Local Similarity 17.9%; Pred. No. 11;  
Matches 34; Conservative 29; Mismatches 54; Indels 73; Gaps 8;

QY 4 VMADEANGEHODAMKHLQKVEENYDLIKATYKNDPVWGNDPTCYG--- 48  
DB 275 LMGSEA-----EDLKRKENKVDLGVNYHHSRVOEPYSSDSLADWRPKYYA 325  
QY 49 -----VGTAAONLNEDKNEAMFPMNADTVYQHTPEKATPD 87  
DB 326 SFNKGVRNADNGMEIHFOIYEIKRQIEDGNI-SWFISEN----- 368  
QY 88 KMYGKKNATYYQTEGDGOVLNDV-LAFSDNCYVYALGPSGSA-GYEIMATDVTDP 145  
DB 369 ---GKGVENEHFKDEGQIODDYRIITTEHLEHLKAIQESNCGYHW-----TP 419  
QY 146 ASCLEKFNEX 155  
DB 420 IDQWNRNSY 429

RESULT 11  
A60272  
Iga-specific metalloendopeptidase (EC 3.4.24.13) - Streptococcus sanguis

N:Alternate names: Iga protease; immunoglobulin A1 proteinase  
C:Species: Streptococcus sanguis  
C>Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #text\_change 21-Jan-2000  
C:Accession: A60272  
R:Gilbert, J.V.; Platt, A.G.; Wright, A.  
Infect. Immun. 59, 7-17, 1991  
A:Title: Analysis of the immunoglobulin A protease gene of Streptococcus sanguis.  
A:Reference number: A60272; MUID:91100011; PMID:1987065  
A:Accession: A60272  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-1668 <GII>  
C:Comment: This protein is shown from the start of translation of this gene as determ  
the start codon is shown in entry B60272.  
C:Genetics:  
A:Gene: Iga  
A:Start codon: GTG

Query Match 8.6%; Score 82; DB 2; Length 1668;  
Best Local Similarity 27.1%; Pred. No. 61;  
Matches 51; Conservative 12; Mismatches 47; Indels 78; Gaps 12;

QY 7 DEANGEHODAMKHLQKVEENYDLIKATYKNDPVWGNDPTCYG--- 1668  
DB 1160 DRYRNNHR-AGALNKFEVDAQETAKRORHRYLDROGREKLYRNILVYDAK 1218  
QY 38 NDEPWGNDPTCYGTAQONLNEDKNEAMFPMNNA-----DYYQHTPEKATPD 87  
DB 1219 ---FDDDTY-----DKATYEAQDSSNPMPKFFGVGKGVYHNKGAATGQ 1264  
QY 88 KMY-GY---KNEKNATY---QTEGDGOVLNDV-LAFSDNCYVYALGPSG--SGAGYELM 137  
DB 1265 SVYVMGYRMDKDGATYTHEMTHD-----SDNEIY---LGGVGRSGGLPPEF 1310  
QY 138 ATTYTDP 145  
DB 1311 AKGLQAP 1318

RESULT 12  
G81933  
hypothetical protein MMA0876 [imported] - Neisseria meningitidis (strain Z2491 serogr  
C:Species: Neisseria meningitidis  
C>Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
C:Accession: G81933  
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo  
Nature 404, 502-506, 2000  
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491  
A:Reference number: A81775; MUID:2022556; PMID:10761919  
A:Accession: G81933  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-133 <PAR>  
A:Cross-references: GB:AL162754; GB:AL157959; NID:97379424; PIDN:CAB84156.1; PID:9737  
A:Experimental source: serogroup A, strain Z2491  
C:Genetics:  
A:Gene: MMA0876

Query Match 8.6%; Score 81.5; DB 2; Length 133;  
Best Local Similarity 24.8%; Pred. No. 2.6;  
Matches 41; Conservative 19; Mismatches 52; Indels 53; Gaps 10;

QY 13 EHODAMKHL-----QKVEENYDLIKATYKNDPVWG---NDFTCVGTAQONLN-EDKNE 64  
DB 3 QHIEKQHLHSDREOKLAE-----VWGLVQNDDEVYHMKLKNAPPEASE 49  
QY 65 AMFEMNADTVYQHTPEKATPDKMYGKKNATYYQTEGDGOVLNDV-LAFS---DNCYV 121  
DB 50 FWRMAETLSTL-----PPNR-----SLDRMNGRLATVAVSLTSMIEDN--- 90





[illegible]

```

OY      141 YTDVP 145
Db      1544 LQAP 1548

RESULT 7
O9L2X9 PRELIMINARY; PRT; 648 AA.
ID 09L2X9
AC 09L2X9;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical protein.
GN T4C21_260 OR A3G60850.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Chojeane N., Robert C., Brotier P., Mincker P., Catolico L.,
RA Attienave F., Saurin W., Weisenbach J., Mewes H.W., Rudd S.,
RA Lemcke K., Mayer K.F.X., Quelet F., Salanoubat M.,
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Southwick A., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,
RA Carninci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,
RA Deng J.M., Hayashizaki Y., Hsuan V.M., Lee J.M., Ishida J., Kamiya A.,
RA Kato J., Kim C.J., Narusaka M., Onodera C.S., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Shinozaki K., Toriumi M., Wong C., Wu H.C.,
RA Yamada K., Yu G., Yuan S., Shinozaki K., Ecker J., Theologis A.,
RA Davis R.M.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL162295; CAB82689.1;
DR EMBL: BF002453; AAC00813.1;
KM Hypothetical protein.
SO SEQUENCE 648 AA; 72914 MW; 83A3F7551172FC58 CRC64;

Query Match
Best Local Similarity 9.2%; Score 88; DB 10; Length 648;
Matches 38; Conservative 19; Mismatches 69; Indels 24; Gaps 6;

OY      2 KPYWADDEANGEHODAMKHLQKLVEEYDIKATYKNDPYWGDFTCVGTAQNLNEDK 61
Db      146 RAWMKRNETATVLTAMKRIOD-----KLTSLRDPATGNEFLCF--KNNSOQFVS 193
OY      62 NVEAM--FMENMADTVYQHTPEKATPDKMYGKNEKATYQEDGCVLTDLAFSDNC 119
Db      194 HINQKODIYVGFADGDKLHSGKRTIDRIKQVWYTGAKLYGIPESIRCVAA-----C 248
OY      120 YVIVALGPDGSGAGYELMATDYT--DVPA 146
Db      249 PVCNA--DSGSASRNKRREPTESLDVPA 276

RESULT 8
O8L2E8 PRELIMINARY; PRT; 677 AA.
ID 08L2E8
AC 08L2E8;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Zinc metalloprotease Pap6.
OS Vibrio Harveyi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OX NCBI_TaxID=669;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RA Teo J., Poh C.L., Zhang L.H.;
RT "Vibrio harveyi zinc metalloprotease.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF508306; AA34261.1;
DR InterPro: IPR001570; Peptidase_M4.
DR InterPro: IPR005075; Pep_M4_propep.
DR InterPro: IPR006025; Zn_MTPeptidase.
DR Pfam: PF01447; Peptidase_M4; 1.
DR Pfam: PF02668; Peptidase_M4_C; 1.
DR Pfam: PF03413; Pep_M4_propep; 1.
DR PRINTS: PR00730; THERMOLYSIN.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
KM Protease; Metalloprotease.
SO SEQUENCE 677 AA; 75120 MW; 5E904C0A127CA186 CRC64;

Query Match
Best Local Similarity 9.2%; Score 88; DB 2; Length 677;
Matches 47; Conservative 26; Mismatches 60; Indels 84; Gaps 10;

OY      2 KPYWADDEANGEHODAMKHLQKLVEE-----NYDL 31
Db      169 RPTLVDAKSGELMDMEGLAFLEAGPGGNKSGRYTPNKKFGFQVDYCCMNSDN 228
OY      32 IKATYKNDPYWGD--FTCVGTAQNLNEDKNEVAMFMENMADTVYQHTF----- 81
Db      229 VETINNNNOOMGQVHRFNC-----NVN-NYREINGAVAPMNDAHYFGQVDFMYREWL 281
OY      82 -----EKATPDKMYGKNEKATYQEDGCVLT-----DYLA-----PS 115
Db      282 GARPIQOKLTVRHYSYNGNAFW---DGRITFGDGNSSMYPLATRDVIAHEVSHGFT 337
OY      116 DDNCYVIVALGPDGSGAGYELMATDYTPVASCLERF 152
Db      338 EQNSGLELR---GMSGWYE-----SFSDVAAALSLQY 367

RESULT 9
O92CG8 PRELIMINARY; PRT; 1237 AA.
ID 092CG8
AC 092CG8;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical protein lin1204.
GN LIN1204.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RA SPRAIN-CLIP 11262 / Serovar 6a;
RX PubMed=11679669;
RA Glaeser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetoui F., Couve E., de Daruvar A., Dehoux P.,
RA Domant E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Esli H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapat G.,
RA Madueno E., Maitounam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstok G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:845-853(2001);
DR EMBL: AL596168; CAC96435.1;
DR Listlist: LIN01204;
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR006192; LPXTG.
DR InterPro: IPR001611; LRR.
DR Pfam: PF00746; Gram_pos_anchor; 1.

```



ID P73122 PRELIMINARY: PRT: 368 AA.  
AC P73122:  
DT 01-FEB-1997 (TREMBLrel. 02, Created)  
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Hypothetical protein slr1920.  
GN SLR1920.  
OS *Synechocystis* sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; *Synechocystis*.  
NCBI\_TaxID=1148;  
OX [1]  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97061201; PubMed=8905231;  
RA Kaneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hirosewa M., Sugitani M., Sasamoto S., Kimura T.,  
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,  
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
RA Tabata S.;  
RT -Sequence analysis of the genome of the unicellular cyanobacterium  
RT *Synechocystis* sp. strain PCC6803. II. Sequence determination of the  
RT entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:109-136(1996).  
DR EMBL: D90903; BAA17148.1;  
KW Hypothetical protein, complete proteome.  
SQ SEQUENCE 368 AA: 43031 MW: 654612A3BDD17F3 CRC64;

Query Match 9.1%; Score 86.5; DB 16; Length 368;  
Best Local Similarity 26.2%; Pred. No. 15;  
Matches 43; Conservative 24; Mismatches 50; Indels 47; Gaps 10;

QY 22 OKLVEENYDL-----IKATYKNDPVNGNDFCTGTAONLNDEKVEAMFMFMNNA 73  
DB 112 OKLYRESFNLQPGHLLVAIKIRKKSIVGFKDY-----ADNLDK-----FLRIFN- 156  
QY 74 DTVYQ---HTFEKATPPDKMVGYNKNATTTQTEGQVLTDLVAFSDDCVYIYAGPFG 129  
DB 157 PAIHOLLIRYFIPVDDMHOPDN---TELPQFEKILTDVINEFDLN---YSL----- 205  
QY 130 SGAGYELW-----ATDYTPVAPASCLEFNEFYAAGLFPVDVYTS 168  
DB 206 DNLTYDKQALVHALDETGVSLNLVKNHGEY-----PFRMEPTFD 245

RESULT 13  
O8IL135 PRELIMINARY: PRT: 1898 AA.  
AC O8IL135:  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Hypothetical protein, conserved.  
GN PF14.0414.  
OS *Plasmodium falciparum* (isolate 3D7).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; *Plasmodium*.  
NCBI\_TaxID=36329;  
OX [1]  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22255705; PubMed=12368864;  
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,  
RA Carlson J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,  
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyte S.,  
RA Chen M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,  
RA Pereira M., Allen J., Selengut J., Haft D., Mather M.W., Valdivia A.B.,  
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,  
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,  
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,  
RA Fraser C.M., Barrell B.;  
RT -Genome sequence of the human malaria parasite *Plasmodium*  
RT *falciparum*.  
RL Nature 419:498-511(2002).  
DR EMBL: AE014823; AAN37027.1; -  
KW Hypothetical protein.

SQ SEQUENCE 1898 AA: 223829 MW: 7B5CB02A7091822F CRC64;

Query Match 9.0%; Score 86; DB 5; Length 1898;  
Best Local Similarity 27.4%; Pred. No. 14e+02;  
Matches 31; Conservative 21; Mismatches 47; Indels 14; Gaps 6;

QY 17 ANKHLOKVEENYDLIKATYKNDP---VGNDFCTGTAONLNDEKVEAMFMFMNNA 73  
DB 571 AMKNSSVLVEE---YCKSKYENKKOKLLIDYDNLSTFTIHYISNNHDSKEQSNFFPHNNA 627  
QY 74 DTVYQHTFEKATPPDKM---YGYNK--ENATTTQTEGQVLTDLVAFSDDCVYI 121  
DB 628 NSSY-VTFDITDISENFISKYFPRISKRSYDTEGEKFDNLIL--GGMCYI 677

RESULT 14  
O9UOH2 PRELIMINARY: PRT: 1928 AA.  
AC O9UOH2:  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE Hypothetical 229.7 kDa protein.  
GN MAL4P2.50.  
OS *Plasmodium falciparum* (isolate 3D7).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; *Plasmodium*.  
NCBI\_TaxID=36329;  
OX [1]  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=3D7;  
RA Devlin K., Bowman S., Churcher C., Harris B., Harris D., Lawson D.,  
RA Quail M., Barrell B.;  
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL035475; CAB62891.1; -  
KW Hypothetical protein  
SQ SEQUENCE 1928 AA: 229692 MW: 38F6CECA9B3253B CRC64;

Query Match 9.0%; Score 86; DB 5; Length 1928;  
Best Local Similarity 21.4%; Pred. No. 1.4e+02;  
Matches 33; Conservative 24; Mismatches 53; Indels 44; Gaps 6;

QY 13 EHODAMKHLOKVEENYDLIKATY-----KNPVGNDP--TVGTAONLNDE 59  
DB 702 DHVQMDKHSIAIINDIDINDVGTSCDNPFPPIKNDRVYDKFEYCNDELFYENIQNV 761  
QY 60 EKNV--EAMFME--MNAADTVYQHTFEKATPPDKMVGYNKNAT----- 98  
DB 762 DONITRKDEFFFNIDNNDHEITCENFSIYKSLNEXEKIKNDKLIELSDDIYSLE 821  
QY 99 TQTEGQVLTDLVAFSDDCVYI 122  
DB 822 SYEGSEGVYIHNKIKISDDDTFSLGMMENNCYII 855

RESULT 15  
O33763 PRELIMINARY: PRT: 1874 AA.  
AC O33763:  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE 1991 protease (EC 3.4.24.13).  
GN IGA.  
OS *Streptococcus sanguis*.  
OC Bacteria; Firmicutes; Lactobacillales; *Streptococcaceae*.  
OC *Streptococcus*.  
NCBI\_TaxID=1305;  
OX [1]  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=SK4;  
RX MEDLINE=98084472; PubMed=9423856;  
RA Poulsen K., Reinholdt J., Jespersgaard C., Boye K., Brown T.A.,  
RA Hauge M., Kilian M.;

DR Pfam; PF00560; IRR; 4.  
DR PRINTS; PRO0019; LEURICHRP.  
DR TIGRFAMS; TIGR01167; LPXTG\_anchor; 1.  
DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 1237 AA; 137104 MW; FDF6556017FB627 CRC64;

Query Match 9.2%; Score 88; DB 16; Length 1237;  
Best Local Similarity 24.8%; Pred. No. 53;  
Matches 41; Conservative 23; Mismatches 65; Indels 36; Gaps 6;

QY 19 KHLQKLVENYDLIKATYK-----DPYMGNDIT-----CVGTAQNLIN 57  
DB 775 KYTVTLNENLDGKAPKREVIYIDAVOGANITVYKEDSGNKLAEISLTGNVEEYS 834  
QY 58 EDEKNVEAMFPMNADIVYQHTFEKATPDKMGYN---KENAITYTEGQYL--TDV 111  
DB 835 SSAKEILGTYLTTEMPNAGEFSLEQVTVYISKNPVPARDIVYQYDEDEGMELAPDET 894  
QY 112 LAFSDNDCYVYIALGPDGSGAGYELMATDYTDVPASCLEKFNEXA 156  
DB 895 LSGVNDENVYTTA-----KITFTGYEL-----IEFSPNAGKFSENA 930

RESULT 10  
033764 PRELIMINARY; PRT; 1854 AA.  
ID 033764  
AC 033764;  
DT 01-JAN-1998 (TREMblrel. 05, Created)  
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)  
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
DE Iga1 protease (EC 3.4.24.13).  
GN IGA.  
OS Streptococcus sanguis.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1305;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SK49;  
RX MEDLINE=98084472; PubMed=9423856;  
RA Poulsen K., Reinholdt J., Jespersgaard C., Boye K., Brown T.A.,  
Hauge M., Kilian M.;  
RT "A comprehensive genetic study of streptococcal immunoglobulin A1  
proteases: evidence for recombination within and between species.";  
RL Infect. Immun. 66:181-190(1998)  
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY  
AN AMIDE BOND (BY SIMILARITY).  
CC EMBL; Y13460; CAAT3857.1; -;  
DR MEROPS; M26.001; -;  
DR InterPro; IPR005877; Gpos\_YsIRK.  
DR InterPro; IPR001899; Gram\_pos\_anchor.  
DR InterPro; IPR006192; LPXTG.  
DR InterPro; IPR006025; Zn\_MTPeptase.  
DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
DR Pfam; PF04650; YsIRK\_signal; 1.  
DR TIGRFAMS; TIGR01167; LPXTG\_anchor; 1.  
DR TIGRFAMS; TIGR01168; YsIRK\_signal; 1.  
DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
KW Cell wall; Hydrolyase; Peptidoglycan-anchor; Protease.  
SQ SEQUENCE 1854 AA; 205915 MW; 09AF960055C8EE30 CRC64;

Query Match 9.2%; Score 88; DB 2; Length 1854;  
Best Local Similarity 27.6%; Pred. No. 89;  
Matches 51; Conservative 13; Mismatches 49; Indels 72; Gaps 13;

QY 7 DEANGEHODAMKHLQKLVEN-----TD-----LIKATYK 37  
DB 1370 DRYANNENHR-AGALINKFVEDNNAOETAKRQCHDHYWRIILDEGKREKLNILLYDAIK 1428  
QY 38 NDPYWGNDFTCVGTAQNLNEDKKNVAMF-----MNNADTVYQHTFEKATPDPMY 90  
P73122

DB 1429 ----FGDD-TTVGKATAEAQFDSSNPAMKYFFGPGVGNKVHNKGAY-----ATGDSVY 1477  
QY 91 --GY--NKENAITY---QTEDEGVLTVDYLAFSDDNDCYVIALGPDG--SGAGYELMATD 140  
DB 1478 YMGYRMLDKGALITYTHEMTD-----SDNETY---LGGYGRSSGLGPEFPKAG 1523  
QY 141 YTDVP 145  
DB 1524 LLOAP 1528

RESULT 11  
033759 PRELIMINARY; PRT; 1854 AA.  
ID 033759  
AC 033759;  
DT 01-JAN-1998 (TREMblrel. 05, Created)  
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)  
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
DE Iga1 protease (EC 3.4.24.13).  
GN IGA.  
OS Streptococcus sanguis.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1305;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SK112;  
RX MEDLINE=98084472; PubMed=9423856;  
RA Poulsen K., Reinholdt J., Jespersgaard C., Boye K., Brown T.A.,  
Hauge M., Kilian M.;  
RT "A comprehensive genetic study of streptococcal immunoglobulin A1  
proteases: evidence for recombination within and between species.";  
RL Infect. Immun. 66:181-190(1998)  
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY  
AN AMIDE BOND (BY SIMILARITY).  
CC EMBL; Y13455; CAAT3852.1; -;  
DR MEROPS; M26.001; -;  
DR InterPro; IPR005877; Gpos\_YsIRK.  
DR InterPro; IPR001899; Gram\_pos\_anchor.  
DR InterPro; IPR006192; LPXTG.  
DR InterPro; IPR006025; Zn\_MTPeptase.  
DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
DR Pfam; PF04650; YsIRK\_signal; 1.  
DR TIGRFAMS; TIGR01167; LPXTG\_anchor; 1.  
DR TIGRFAMS; TIGR01168; YsIRK\_signal; 1.  
DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
KW Cell wall; Hydrolyase; Peptidoglycan-anchor; Protease.  
SQ SEQUENCE 1854 AA; 206031 MW; 0538AB0389968B58 CRC64;

Query Match 9.2%; Score 88; DB 2; Length 1854;  
Best Local Similarity 27.6%; Pred. No. 89;  
Matches 51; Conservative 13; Mismatches 49; Indels 72; Gaps 13;

QY 7 DEANGEHODAMKHLQKLVEN-----TD-----LIKATYK 37  
DB 1370 DRYANNENHR-AGALINKFVEDNNAOETAKRQCHDHYWRIILDEGKREKLNILLYDAIK 1428  
QY 38 NDPYWGNDFTCVGTAQNLNEDKKNVAMF-----MNNADTVYQHTFEKATPDPMY 90  
DB 1429 ----FGDD-TTVGKATAEAQFDSSNPAMKYFFGPGVGNKVHNKGAY-----ATGDSVY 1477  
QY 91 --GY--NKENAITY---QTEDEGVLTVDYLAFSDDNDCYVIALGPDG--SGAGYELMATD 140  
DB 1478 YMGYRMLDKGALITYTHEMTD-----SDNETY---LGGYGRSSGLGPEFPKAG 1523  
QY 141 YTDVP 145  
DB 1524 LLOAP 1528

RESULT 12  
P73122

RT "A comprehensive genetic study of streptococcal immunoglobulin A1  
RT proteases: evidence for recombination within and between species.";  
RL Infect. Immun. 66:181-190(1998).  
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY  
CC AN AMIDE BOND (BY SIMILARITY).  
CC EMBL: Y13459; CAA73856.1; -.  
DR MEROPS: M26.001; -.  
DR InterPro: IPR005877; Gpos\_YsIRK.  
DR InterPro: IPR001899; Gram\_pos\_anchor.  
DR InterPro: IPR006192; LPXTG.  
DR InterPro: IPR006025; Zn\_MTpeptidse.  
DR Pfam: PF00746; Gram\_pos\_anchor; 1.  
DR Pfam: PF04650; YsIRK\_signal; 1.  
DR TIGRFAMs: TIGR01167; LPXTG\_anchor; 1.  
DR TIGRFAMs: TIGR01168; YsIRK\_signal; 1.  
DR PROSITE: PS50847; Gram\_pos\_anchoring; 1.  
DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
KW Cell wall; Hydrolase; Peptidoglycan-anchor; Protease.  
SO SEQUENCE 1874 AA; 207888 MW; FA30684926B84532 CRC64;

Query Match 8.98; Score 85; DB 2; Length 1874;

Best Local Similarity 27.68; Pred. No. 1.7e+02;  
Matches 51; Conservative 13; Mismatches 49; Indels 72; Gaps 13;

```
OY 7 DEANGSEHODAMKHLQKLVEN-----YD-----LIKATYK 37
DB 1390 DRYRNNNEHK-AGAELEKFEVDNAQETAKRQRDHDYRYRIIDEGREKLYRNILLYDAYK 1448
OY 38 NDPVWGNDFTCVGTAQNLNDEKNEVAMFMF-----MNNADTVYQHTFEKATPDPMY 90
DB 1449 ----FGDDAT-VGKATAEAQFDSSNPAMKYEFGVGNKVVHNRGAY-----ATGDGY 1497
OY 91 --GY--NKENATY---QTEDEQVLTDLVAFSPDNCYVYIALGPDG--SGAGYELMATD 140
DB 1498 YMGYRMLDKDGAITYTHEMTHD-----SDNETY---LGGYGRSRGLGPEFFPAKG 1543
OY 141 YTDVP 145
DB 1544 LLOAP 1548
```

Search completed: August 1, 2003, 12:19:57  
Job time : 51.4705 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: August 1, 2003, 12:15:58 ; Search time 18.0267 Seconds

(without alignments)  
1199.017 Million cell updates/sec

Title: US-10-087-195-8

Perfect score: 989  
Sequence: 1 NPTMANEAKLGSYODAMKSL.....EKTVRNYTDSCKPAPAQN 182

## Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

## Database :

Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpae/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpae/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpae/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpae/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpae/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpae/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpae/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpae/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpae/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpae/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpae/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpae/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpae/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpae/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpae/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpae/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpae/US60\_PUBCOMB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpae/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	989	100.0	182	US-10-085-572-8	Sequence 8, Appl1
2	989	100.0	182	US-10-087-195-8	Sequence 8, Appl1
3	331.5	33.5	171	US-10-085-572-7	Sequence 7, Appl1
4	331.5	33.5	171	US-10-087-195-7	Sequence 7, Appl1
5	267.5	27.0	172	US-10-085-572-6	Sequence 6, Appl1
6	267.5	27.0	172	US-10-087-195-6	Sequence 6, Appl1
7	83.5	8.4	778	US-09-815-242-10728	Sequence 10728, A
8	79	8.0	221	US-09-728-914-4	Sequence 4, Appl1
9	76	7.7	2201	US-09-995-542-9	Sequence 9, Appl1
10	74.5	7.5	142	US-09-893-737-172	Sequence 172, App
11	74.5	7.5	198	US-09-864-761-48672	Sequence 48672, A
12	74.5	7.5	199	US-09-801-231-2	Sequence 2, Appl1
13	74.5	7.5	199	US-09-731-872-269	Sequence 269, App
14	74.5	7.5	199	US-10-052-586-588	Sequence 588, App
15	74.5	7.5	199	US-10-174-590-588	Sequence 588, App

16	74.5	7.5	199	15	US-10-176-758-588	Sequence 588, App
17	74.5	7.5	199	15	US-10-175-737-588	Sequence 588, App
18	74.5	7.5	199	15	US-10-173-706-588	Sequence 588, App
19	74.5	7.5	199	15	US-10-175-738-588	Sequence 588, App
20	74.5	7.5	199	15	US-10-175-752-588	Sequence 588, App
21	74.5	7.5	199	15	US-10-176-482-588	Sequence 588, App
22	74.5	7.5	199	15	US-10-176-757-588	Sequence 588, App
23	74.5	7.5	199	15	US-10-176-913-588	Sequence 588, App
24	74.5	7.5	199	15	US-10-180-532-588	Sequence 588, App
25	74.5	7.5	199	15	US-10-180-557-588	Sequence 588, App
26	74.5	7.5	199	15	US-10-173-700-588	Sequence 588, App
27	74.5	7.5	199	15	US-10-174-572-588	Sequence 588, App
28	74.5	7.5	199	15	US-10-174-579-588	Sequence 588, App
29	74.5	7.5	199	15	US-10-174-582-588	Sequence 588, App
30	74.5	7.5	199	15	US-10-174-588-588	Sequence 588, App
31	74.5	7.5	199	15	US-10-175-739-588	Sequence 588, App
32	74.5	7.5	199	15	US-10-175-740-588	Sequence 588, App
33	74.5	7.5	199	15	US-10-175-743-588	Sequence 588, App
34	74.5	7.5	199	15	US-10-176-488-588	Sequence 588, App
35	74.5	7.5	199	15	US-10-176-492-588	Sequence 588, App
36	74.5	7.5	199	15	US-10-176-747-588	Sequence 588, App
37	74.5	7.5	199	15	US-10-176-750-588	Sequence 588, App
38	74.5	7.5	199	15	US-10-176-985-588	Sequence 588, App
39	74.5	7.5	199	15	US-10-176-987-588	Sequence 588, App
40	74.5	7.5	199	15	US-10-176-992-588	Sequence 588, App
41	74.5	7.5	199	15	US-10-176-993-588	Sequence 588, App
42	74.5	7.5	199	15	US-10-184-658-588	Sequence 588, App
43	74.5	7.5	199	15	US-10-176-991-588	Sequence 588, App
44	74.5	7.5	199	15	US-10-173-695-588	Sequence 588, App
45	74.5	7.5	199	15	US-10-173-697-588	Sequence 588, App

## ALIGNMENTS

RESULT 1  
US-10-085-572-8  
Sequence 8, Application US/10085572  
Publication No. US20020151499A1  
GENERAL INFORMATION:  
APPLICANT: Nuttall, Patricia, Ann  
APPLICANT: Paesen, Guido, Christiaan  
TITLE OF INVENTION: Treatment of Conjunctionitis  
FILE REFERENCE: 2488-1-003  
CURRENT APPLICATION NUMBER: US/10/085,572  
CURRENT FILING DATE: 2002-02-27  
PRIOR APPLICATION NUMBER: PCT/GB00/03282  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: 9920674.0  
PRIOR FILING DATE: 1999-09-01  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8  
LENGTH: 182  
TYPE: PRT  
ORGANISM: MS-HBP 1  
US-10-085-572-8

Query Match	Score	DB ID	Length
Best Local Similarity	100.0%	Pred. No. 3,7e-99;	182;
Matches 182; Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;
1 NPTMANEAKLGSYODAMKSL	100.0%	US-10-085-572-8	182
1 NPTMANEAKLGSYODAMKSL	100.0%	US-10-087-195-8	182
61 NATIYKXKHLTLKESHETITVWKAYDTTENGKIKYETGRTOTFFEDVEFYFSYKNC	100.0%	US-10-085-572-7	171
61 NATIYKXKHLTLKESHETITVWKAYDTTENGKIKYETGRTOTFFEDVEFYFSYKNC	100.0%	US-10-087-195-7	171
121 VIFVPEKRGSDGDEYELWVSEDKIDKIDPCKFTMAVPAQOQEKTVRNVYDSSCKPAPA	100.0%	US-10-085-572-6	172
121 VIFVPEKRGSDGDEYELWVSEDKIDKIDPCKFTMAVPAQOQEKTVRNVYDSSCKPAPA	100.0%	US-10-087-195-6	172
121 VIFVPEKRGSDGDEYELWVSEDKIDKIDPCKFTMAVPAQOQEKTVRNVYDSSCKPAPA	100.0%	US-09-815-242-10728	778
121 VIFVPEKRGSDGDEYELWVSEDKIDKIDPCKFTMAVPAQOQEKTVRNVYDSSCKPAPA	100.0%	US-09-728-914-4	221
121 VIFVPEKRGSDGDEYELWVSEDKIDKIDPCKFTMAVPAQOQEKTVRNVYDSSCKPAPA	100.0%	US-09-995-542-9	2201
61 NATIYKXKHLTLKESHETITVWKAYDTTENGKIKYETGRTOTFFEDVEFYFSYKNC	100.0%	US-09-893-737-172	142
61 NATIYKXKHLTLKESHETITVWKAYDTTENGKIKYETGRTOTFFEDVEFYFSYKNC	100.0%	US-09-864-761-48672	198
121 VIFVPEKRGSDGDEYELWVSEDKIDKIDPCKFTMAVPAQOQEKTVRNVYDSSCKPAPA	100.0%	US-09-801-231-2	199
121 VIFVPEKRGSDGDEYELWVSEDKIDKIDPCKFTMAVPAQOQEKTVRNVYDSSCKPAPA	100.0%	US-09-731-872-269	199
121 VIFVPEKRGSDGDEYELWVSEDKIDKIDPCKFTMAVPAQOQEKTVRNVYDSSCKPAPA	100.0%	US-10-052-586-588	199
121 VIFVPEKRGSDGDEYELWVSEDKIDKIDPCKFTMAVPAQOQEKTVRNVYDSSCKPAPA	100.0%	US-10-174-590-588	199

Qy 181 QN 182  
Db 181 QN 182

## RESULT 2

US-10-087-195-8  
; Sequence 8, Application US/10087195  
; Publication No. US20020193306A1  
; GENERAL INFORMATION:  
; APPLICANT: Nuttall, Patricia, Anne  
; TITLE OF INVENTION: Treatment of Allergic Rhinitis  
; FILE REFERENCE: 2488-1-004  
; CURRENT APPLICATION NUMBER: US/10/087,195  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: PCT/GB00/03287  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 9920673.2  
; PRIOR FILING DATE: 1999-09-01  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 182  
; TYPE: PRT  
; ORGANISM: MS-HBP 1  
US-10-087-195-8

Query Match 100.0%; Score 989; DB 14; Length 182;  
Best Local Similarity 100.0%; Pred. No. 3.7e-99;  
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NPTWANEAKLSGYODAMKSLQODONKRRYLAQATQTTDGVGGEFTCVSYAEKIGKKL 60  
Db 1 NPTWANEAKLSGYODAMKSLQODONKRRYLAQATQTTDGVGGEFTCVSYAEKIGKKL 60  
Qy 61 NATILYKKNKHLTDKESHEITVWKAYDYTTENGIRKETGCTRTQTFEDVFPVSDYKNC 120  
Db 61 NATILYKKNKHLTDKESHEITVWKAYDYTTENGIRKETGCTRTQTFEDVFPVSDYKNC 120  
Qy 121 VIFVPEKRGSDDEGYELWVSEDKIDKIPDCCKFTMAVFAQOQEKTVRNYTTDSSCKPAPA 180  
Db 121 VIFVPEKRGSDDEGYELWVSEDKIDKIPDCCKFTMAVFAQOQEKTVRNYTTDSSCKPAPA 180  
Qy 181 QN 182  
Db 181 QN 182

## RESULT 3

US-10-085-572-7  
; Sequence 7, Application US/10085572  
; Publication No. US20020151499A1  
; GENERAL INFORMATION:  
; APPLICANT: Nuttall, Patricia, Ann  
; APPLICANT: Paesen, Guido, Christiaan  
; TITLE OF INVENTION: Treatment of Conjunctivitis  
; FILE REFERENCE: 2488-1-003  
; CURRENT APPLICATION NUMBER: US/10/085,572  
; PRIOR FILING DATE: 2002-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB00/03282  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 9920674.0  
; PRIOR FILING DATE: 1999-09-01  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 171  
; TYPE: PRT  
; ORGANISM: FS-HBP 2  
US-10-085-572-7

Query Match 33.5%; Score 331.5; DB 14; Length 171;  
Best Local Similarity 42.0%; Pred. No. 5.3e-28;  
Matches 74; Conservative 29; Mismatches 62; Indels 11; Gaps 7;

Qy 2 PTWANEAKLSGYODAMKSLQODONKRRYLAQATQTTDGVGGEFTCVSYAEKIG--KKK 59  
Db 3 PDMADEAANGAHQADAMKSLKADVENYVWKATYKNPQWQNDFTCVGAMANDVNEDEKS 62  
Qy 60 LNATILYKKNKHLTDKESHEITVWKAYDYTTENGIRKETGCTRTQTFEDVFPVSDYKNC 119  
Db 63 IQAEFLFMNNADTNMOPATEKVTAVKMYGYNRENAFRETED--GQVFTDVIAYSD--DNC 119  
Qy 120 DVIVPEKRGSDDEGYELWVSEDKIDKIPDCCKFTMAVFAQOQEKTVRNYTTDSSC 175  
Db 120 DVIVPEKRGSDDEGYELWVSEDKIDKIPDCCKFTMAVFAQOQEKTVRNYTTDSSC 175

## RESULT 4

US-10-087-195-7  
; Sequence 7, Application US/10087195  
; Publication No. US20020193306A1  
; GENERAL INFORMATION:  
; APPLICANT: Nuttall, Patricia, Anne  
; TITLE OF INVENTION: Treatment of Allergic Rhinitis  
; FILE REFERENCE: 2488-1-004  
; CURRENT APPLICATION NUMBER: US/10/087,195  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: PCT/GB00/03287  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 9920673.2  
; PRIOR FILING DATE: 1999-09-01  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 171  
; TYPE: PRT  
; ORGANISM: FS-HBP 2  
US-10-087-195-7

Query Match 33.5%; Score 331.5; DB 14; Length 171;  
Best Local Similarity 42.0%; Pred. No. 5.3e-28;  
Matches 74; Conservative 29; Mismatches 62; Indels 11; Gaps 7;

Qy 2 PTWANEAKLSGYODAMKSLQODONKRRYLAQATQTTDGVGGEFTCVSYAEKIG--KKK 59  
Db 3 PDMADEAANGAHQADAMKSLKADVENYVWKATYKNPQWQNDFTCVGAMANDVNEDEKS 62  
Qy 60 LNATILYKKNKHLTDKESHEITVWKAYDYTTENGIRKETGCTRTQTFEDVFPVSDYKNC 119  
Db 63 IQAEFLFMNNADTNMOPATEKVTAVKMYGYNRENAFRETED--GQVFTDVIAYSD--DNC 119  
Qy 120 DVIVPEKRGSDDEGYELWVSEDKIDKIPDCCKFTMAVFAQOQEKTVRNYTTDSSC 175  
Db 120 DVIVPEKRGSDDEGYELWVSEDKIDKIPDCCKFTMAVFAQOQEKTVRNYTTDSSC 175

## RESULT 5

US-10-085-572-6  
; Sequence 6, Application US/10085572  
; Publication No. US20020151499A1  
; GENERAL INFORMATION:  
; APPLICANT: Nuttall, Patricia, Ann  
; APPLICANT: Paesen, Guido, Christiaan  
; TITLE OF INVENTION: Treatment of Conjunctivitis  
; FILE REFERENCE: 2488-1-003  
; CURRENT APPLICATION NUMBER: US/10/085,572  
; PRIOR FILING DATE: 2002-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB00/03282  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 9920674.0  
; PRIOR FILING DATE: 1999-09-01  
; NUMBER OF SEQ ID NOS: 8

```

: SOFTWARE: Fastseq for Windows Version 4.0
: SEQ ID NO: 6
: LENGTH: 172
: TYPE: prt
: ORGANISM: FS-HBP1
: US-10-085-572-6

```

Query Match	27.0%;	Score 267.5;	DB 14;	Length 172;
Best Local Similarity	39.3%;	Pred. No. 4.5e-21;		
Matches 70; Conservative	20;	Mismatches 77;	Indels 11;	Gaps 7;

b

```
2 PTWANEAKLSYQDANKSLQDDONKRYLAQTGTDSVWGEETCVSTAEEKG--KKK 59  
   | | : | | | : | | | | | : | | | | | : | : : |  
3 PWADEAANGHQDAWKHLQKLEENYDLIKATYKNPVGWGNDFTCVGTAQNINDEKEN 62
```

[illegible]

120 DVIFVPERGSDGDYELVNSEDIKIPDCKFTMAYFAQGEKTVRNVYTDSSCKP 177  
11: 11 1111 : 1 1 11:111 11  
120 YVYALCPDGGSGAG-YELMAT-DYIDVPASCLEKFNEYAA---GLPVRDYVT-SDCLP 171

RESULT 6  
US-10-087-195-6  
; Sequence 6, Application US/10087195  
; Publication No. US20020193306A1  
Journal: TROUSSE

```

; APPLICANT: Nuttall, Patricia, Anne
;
; APPLICANT: Paesen, Guido, Christiaan
;
; TITLE OF INVENTION: Treatment of Allergic Rhinitis
;
; FIVE DATES: 2400 1 1 004

```

;  
; CURRENT APPLICATION NUMBER: US/10/087,195  
;  
; CURRENT FILING DATE: 2002-03-01  
;  
; PRIOR APPLICATION NUMBER: PCT/GB00/03287  
;  
; PRIOR FILING DATE: 2000-08-24

```

; PRIOR APPLICATION NUMBER: 9920673.2
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Verst

```

```

; SEQ ID NO 6
;
; LENGTH: 172
;
; TYPE: PRT
;
; ORGANISM: FS-HBP1
;

```

Query Match	27.08;	Score 267.5;	DB 14;	Length 172;
Best Local Similarity	39.38;	Pred. No. 4.5e-21;		
Matches 70; Conservative	20;	Mismatches 77;	Indels 11;	Gaps 77;

2 PTWANEAKLGSTODAMKSLQDDQNKRYLAQATQTTDSVWGEEPTCVSTAEGT-KK 59  
 3 PVMADERANGEHQDAMKHLQKLVLEENDLIKRTYKNDVWGNDPFTCVGTAQGNLNDEKN 62

60 LNTILYKKKHLTDLKESHEITTVKAYDYTTENGIKETQGTROTFFEDYVFSKYNC 119  
 : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
63 VEAMFMFNNAADVYQHTFEKATPDKMYGYKNKENAITYQTED--GQLTDLIAFSD-DNC 119

120 DYIFVPRKRGSDGDEYELWNSEDKIDKIPDCCKFTMAIYFAQQQEKTVRNVYTTDSSCKP 177  
 ||: || ||||| : || : ||: ||| |||  
 120 VYIYALGPDGSGAG-YELNWT-DYTDVPAACLEKFNEYAA---GLPVRDVTY-SDCLP 171

RESULT 7  
US-09-815-242-10728  
; Sequence 10728, Application US/09815242  
; Patent No. US20020061569A1

```

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: zyskind, Judith W
; APPLICANT: wall, Daniel

```

APPLICANT: Trawick, John D.  
 APPLICANT: Carr, Grant J.  
 APPLICANT: Yamamoto, Robert T.  
 APPLICANT: Xu, H. Howard  
 TITLE OF INVENTION: Identification of Essential Genes in  
 TITLE OF INVENTION: Prokaryotes

CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIORITY DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/269,308  
 PRIOR FILING DATE: 2001-02-16

```
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10728
LENGTH: 778
TYPE: PRT
```

8.48; S

Matches 43; Conservative 25; Mismatches 62; Indels 45; Gaps 11.

14 QDAWKSLLQDDQNK-----RYVLAQATQTDDGVNGEFTCVSYTAERIGKKKLNATILYK 67

176 QEAAMAVRLEREKSEKEIITYI -NKVYMANGFYGH-----TAEN-----YY 218  
68 NKHLTDLKESHETIVM-----KAYD-YTTENGIKYETQRTOTFEDVFVPSDYKNCV 121

219 GKHLSELDLPGALLAGMGPQAPNSYDPRY-----KBPDAKER-----RDVLELYTMBDKKK  
122 IFVPERGS-----DEGYELWSEDKIDKIPDCKFTMAYFAOQOEKTVRNVYTD 1722

SULT 8  
-09-728-914-4

PATENT NO. US20010046499A1  
GENERAL INFORMATION:  
APPLICANT: KANTOR, FRED S.

APPLICANT: DAS, SUBRATA  
TITLE OF INVENTION: TICK ANTIGENS AND COMPOSITIONS AND METHODS COMPRISING  
TITLE OF INVENTION: THEM

CURRENT APPLICATION NUMBER: US/09/7728,914  
CURRENT FILING DATE: 2000-12-01  
PRIOR APPLICATION NUMBER: 60/169,048

PRIOR FILING DATE: 2000-10-16  
NUMBER OF SEQ ID NOS: 39  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4

TYPE: PRT  
ORGANISM: Ixodes scapularis  
-09-728-914-4

RESULT 11  
US-09-864-761-48672  
: Sequence 48672, Application US/09864761  
: Patent No. US20020048763A1  
: GENERAL INFORMATION:  
: APPLICANT: Penn, Sharon G.  
: APPLICANT: Rank, David R.  
: APPLICANT: Hanzel, David K.  
: APPLICANT: Chen, Wensheng  
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
: FILE REFERENCE: Aecm1ca-x-1  
: CURRENT APPLICATION NUMBER: US/09/864,761  
: CURRENT FILING DATE: 2001-05-23  
: PRIOR APPLICATION NUMBER: US 60/180,312  
: PRIOR FILING DATE: 2000-02-04  
: PRIOR APPLICATION NUMBER: US 60/207,456  
: PRIOR FILING DATE: 2000-05-26  
: PRIOR APPLICATION NUMBER: US 09/632,366  
: PRIOR FILING DATE: 2000-08-03  
: PRIOR APPLICATION NUMBER: GB 24263,6  
: PRIOR FILING DATE: 2000-10-04  
: PRIOR APPLICATION NUMBER: US 60/236,359  
: PRIOR FILING DATE: 2000-09-27  
: PRIOR APPLICATION NUMBER: PCT/US01/00666  
: PRIOR FILING DATE: 2001-01-30  
: PRIOR APPLICATION NUMBER: PCT/US01/00667  
: PRIOR FILING DATE: 2001-01-30  
: PRIOR APPLICATION NUMBER: PCT/US01/00664  
: PRIOR FILING DATE: 2001-01-30  
: PRIOR APPLICATION NUMBER: PCT/US01/00669  
: PRIOR FILING DATE: 2001-01-30  
: PRIOR APPLICATION NUMBER: PCT/US01/00665  
: PRIOR FILING DATE: 2001-01-30  
: PRIOR APPLICATION NUMBER: PCT/US01/00668  
: PRIOR FILING DATE: 2001-01-30  
: PRIOR APPLICATION NUMBER: PCT/US01/00663  
: PRIOR FILING DATE: 2001-01-30  
: PRIOR APPLICATION NUMBER: PCT/US01/00662  
: PRIOR FILING DATE: 2001-01-30  
: PRIOR APPLICATION NUMBER: PCT/US01/00661  
: PRIOR FILING DATE: 2001-01-30  
: PRIOR APPLICATION NUMBER: PCT/US01/00670  
: PRIOR FILING DATE: 2001-01-30

```
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 48672
;; LENGTH: 198
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC010766.2
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
;; OTHER INFORMATION: EST_HUMAN HIT: AW188146.1, EVALU 1.00e-77
;; OTHER INFORMATION: SWISSPROT HIT: P00673, EVALU 1.00e-05
US-09-864-761-48672
```

```
Query Match
Best Local Similarity 7.5%; Score 74.5; DB 9; Length 198;
Matches 33; Conservative 28; Mismatches 51; Indels 47; Gaps 7;
```

```
OY 30 LAQATQTTDGVGGEFTCVSYAE--KIGKKR-----LNATILYKKNHLDLKEH-HET 80
DB 14 LAEASESTMKIKKEFTDEEMQYDMAKSGQKQTEILMNPILLVKNLSLSMSKDDMSST 73
OY 81 ITWKADYTTENGKIKETGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 137
DB 74 LITFRSLHYNDPKG-----NSSGNDKECCNDMTV 102
OY 138 WVEDKIDKIDPDCCKFTMAVFAQQOEKTVRVNVTDSCK 176
DB 103 W---RKVSEANGSCKMS-NNFIIRSTEVMRRVHRAPSCK 137
```

```
RESULT 12
US-09-801-231-2
;; Sequence 2, Application US/09801231
;; Patent No. US20010049434A1
;; GENERAL INFORMATION:
;; APPLICANT: Conklin, Darrell C.
;; TITLE OF INVENTION: Human Ribonuclease
;; FILE REFERENCE: 00-19
;; CURRENT APPLICATION NUMBER: US/09/801,231
;; CURRENT FILING DATE: 2001-03-07
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 2
;; LENGTH: 199
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-801-231-2
```

```
Query Match
Best Local Similarity 7.5%; Score 74.5; DB 9; Length 199;
Matches 33; Conservative 28; Mismatches 51; Indels 47; Gaps 7;
```

```
OY 30 LAQATQTTDGVGGEFTCVSYAE--KIGKKR-----LNATILYKKNHLDLKEH-HET 80
DB 15 LAEASESTMKIKKEFTDEEMQYDMAKSGQKQTEILMNPILLVKNLSLSMSKDDMSST 74
OY 81 ITWKADYTTENGKIKETGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 137
DB 75 LITFRSLHYNDPKG-----NSSGNDKECCNDMTV 103
OY 138 WVEDKIDKIDPDCCKFTMAVFAQQOEKTVRVNVTDSCK 176
DB 104 W---RKVSEANGSCKMS-NNFIIRSTEVMRRVHRAPSCK 137
```

```
RESULT 13
US-09-731-872-269
```

```
;; Sequence 269, Application US/09731872
;; Patent No. US20020102604A1
;; GENERAL INFORMATION:
;; APPLICANT: Dumas Milne Edwards, Jean Baptiste
;; APPLICANT: Bouguetel, Lydie
;; APPLICANT: Jobert, Severin
;; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
;; FILE REFERENCE: 78-US3-RGS
;; CURRENT APPLICATION NUMBER: US/09/731,872
;; CURRENT FILING DATE: 2000-12-07
;; PRIOR APPLICATION NUMBER: US 60/169,629
;; PRIOR FILING DATE: 1999-12-08
;; PRIOR APPLICATION NUMBER: US 60/187,470
;; PRIOR FILING DATE: 2000-03-06
;; NUMBER OF SEQ ID NOS: 482
;; SOFTWARE: Patent.pm
;; SEQ ID NO 269
;; LENGTH: 199
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SIGNAL
;; LOCATION: -16...-1
US-09-731-872-269
```

```
Query Match
Best Local Similarity 7.5%; Score 74.5; DB 10; Length 199;
Matches 33; Conservative 28; Mismatches 51; Indels 47; Gaps 7;
```

```
OY 30 LAQATQTTDGVGGEFTCVSYAE--KIGKKR-----LNATILYKKNHLDLKEH-HET 80
DB 15 LAEASESTMKIKKEFTDEEMQYDMAKSGQKQTEILMNPILLVKNLSLSMSKDDMSST 74
OY 81 ITWKADYTTENGKIKETGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 137
DB 75 LITFRSLHYNDPKG-----NSSGNDKECCNDMTV 103
OY 138 WVEDKIDKIDPDCCKFTMAVFAQQOEKTVRVNVTDSCK 176
DB 104 W---RKVSEANGSCKMS-NNFIIRSTEVMRRVHRAPSCK 137
```

```
RESULT 14
US-10-052-586-588
;; Sequence 588, Application US/10052586
;; Publication No. US20020127584A1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Chen, Jian
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Pan, James
;; APPLICANT: Smith, Victoria
;; APPLICANT: Watanabe, Colin K.
;; APPLICANT: Wood, William I.
;; APPLICANT: Zhang, Zemin
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P3430R1C1
;; CURRENT APPLICATION NUMBER: US/10/052,586
;; CURRENT FILING DATE: 2002-01-15
;; PRIOR APPLICATION NUMBER: 60/059263
;; PRIOR FILING DATE: 1997-09-18
;; PRIOR APPLICATION NUMBER: 60/059266
;; PRIOR FILING DATE: 1997-09-18
;; PRIOR APPLICATION NUMBER: 60/062250
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/063120
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/063121
;; PRIOR FILING DATE: 1997-10-24
```

PRIOR APPLICATION NUMBER: 60/063486	PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540	PRIOR FILING DATE: 1997-10-26
PRIOR APPLICATION NUMBER: 60/063541	PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544	PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544	PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063546	PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063734	PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063870	PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/064103	PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/065311	PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066120	PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066466	PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066772	PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/067745	PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069335	PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069425	PRIOR FILING DATE: 1997-12-12
PRIOR APPLICATION NUMBER: 60/069870	PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/068017	PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/077450	PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632	PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649	PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/078866	PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939	PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079664	PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786	PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080107	PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194	PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327	PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333	PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081049	PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081838	PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081838	PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955	PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568	PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569	PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569	PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704	PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797	PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/083322	PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083455	PRIOR FILING DATE: 1998-04-28

1	PRIOR FILING DATE: 1998-04-29
2	PRIOR APPLICATION NUMBER: 60/083496
3	PRIOR FILING DATE: 1998-04-29
4	PRIOR APPLICATION NUMBER: 60/083499
5	PRIOR FILING DATE: 1998-04-29
6	PRIOR APPLICATION NUMBER: 60/083555
7	PRIOR FILING DATE: 1998-04-29
8	PRIOR APPLICATION NUMBER: 60/083436
9	PRIOR FILING DATE: 1998-05-05
10	PRIOR APPLICATION NUMBER: 60/084414
11	PRIOR FILING DATE: 1998-05-06
12	PRIOR APPLICATION NUMBER: 60/084635
13	PRIOR FILING DATE: 1998-05-07
14	PRIOR APPLICATION NUMBER: 60/084644
15	PRIOR FILING DATE: 1998-05-07
16	PRIOR APPLICATION NUMBER: 60/085573
17	PRIOR FILING DATE: 1998-05-15
18	PRIOR APPLICATION NUMBER: 60/085579
19	PRIOR FILING DATE: 1998-05-15
20	PRIOR APPLICATION NUMBER: 60/085580
21	PRIOR FILING DATE: 1998-05-15
22	PRIOR APPLICATION NUMBER: 60/085582
23	PRIOR FILING DATE: 1998-05-15
24	PRIOR APPLICATION NUMBER: 60/085700
25	PRIOR FILING DATE: 1998-05-15
26	PRIOR APPLICATION NUMBER: 60/086023
27	PRIOR FILING DATE: 1998-05-18
28	PRIOR APPLICATION NUMBER: 60/086339
29	PRIOR FILING DATE: 1998-05-22
30	PRIOR APPLICATION NUMBER: 60/086486
31	PRIOR FILING DATE: 1998-05-22
32	PRIOR APPLICATION NUMBER: 60/087058
33	PRIOR FILING DATE: 1998-05-28
34	PRIOR APPLICATION NUMBER: 60/087208
35	PRIOR FILING DATE: 1998-05-28
36	PRIOR APPLICATION NUMBER: 60/087609
37	PRIOR FILING DATE: 1998-06-02
38	PRIOR APPLICATION NUMBER: 60/087759
39	PRIOR FILING DATE: 1998-06-02
40	PRIOR APPLICATION NUMBER: 60/088029
41	PRIOR FILING DATE: 1998-06-04
42	PRIOR APPLICATION NUMBER: 60/088033
43	PRIOR FILING DATE: 1998-06-04
44	PRIOR APPLICATION NUMBER: 60/088025
45	PRIOR FILING DATE: 1998-06-04
46	PRIOR APPLICATION NUMBER: 60/088028
47	PRIOR FILING DATE: 1998-06-04
48	PRIOR APPLICATION NUMBER: 60/088029
49	PRIOR FILING DATE: 1998-06-04
50	PRIOR APPLICATION NUMBER: 60/088033
51	PRIOR FILING DATE: 1998-06-04
52	PRIOR APPLICATION NUMBER: 60/088167
53	PRIOR FILING DATE: 1998-06-05
54	PRIOR APPLICATION NUMBER: 60/088202
55	PRIOR FILING DATE: 1998-06-05
56	PRIOR APPLICATION NUMBER: 60/088212
57	PRIOR FILING DATE: 1998-06-05
58	PRIOR APPLICATION NUMBER: 60/088217
59	PRIOR FILING DATE: 1998-06-05
60	PRIOR APPLICATION NUMBER: 60/088326
61	PRIOR FILING DATE: 1998-06-04
62	PRIOR APPLICATION NUMBER: 60/088555
63	PRIOR FILING DATE: 1998-06-05
64	PRIOR APPLICATION NUMBER: 60/088722
65	PRIOR FILING DATE: 1998-06-10
66	PRIOR APPLICATION NUMBER: 60/088738
67	PRIOR FILING DATE: 1998-06-10
68	PRIOR APPLICATION NUMBER: 60/088740
69	PRIOR FILING DATE: 1998-06-10
70	PRIOR APPLICATION NUMBER: 60/088811
71	PRIOR FILING DATE: 1998-06-10
72	PRIOR APPLICATION NUMBER: 60/088824
73	PRIOR FILING DATE: 1998-06-10

Query Match	7.5%;	Score 74.5;	DB 14;	Length 199;
Best Local Similarity	20.8%;	Pred. NO. 4.4;		
Matches 33; Conservative	28;	Mismatches 51;	Indels 47;	Gaps 77;

RESULT 15  
US-10-174-590-588  
; Sequence 588, Application US/10174590  
; Publication No. US20030008352A1

```

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3430R1C42
CURRENT APPLICATION NUMBER: US/10/174,590
CURRENT FILING DATE: 2002-06-18
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 588
LENGTH: 199
TYPE: PRF
ORGANISM: Homo Sapien
US-10-174-590-588

```

Query match 7.58; Score 74.5; DB 15; Length 199;

Best Local Similarity 20.88; Pred. No. 4.4;  
Matches 33; Conservative 28; Mismatches 51; Indels 47; Gaps 7.

[illegible]

Search completed: August 1, 2003, 12:23:46  
Job time : 19.0267 secs

THIS PAGE BLANK (USPTO)



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

Run on: August 1, 2003, 12:15:04 ; Search time 13.8667 Seconds  
(without alignments)  
555.330 Million cell updates/sec

Title: US-10-087-195-8  
 Perfect score: 989  
 Sequence: 1 NPTWANEAKLGSYQDAWKSL.....EKTIVRNVTYTDSSCKPAPAQN 1822

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%

```
Database :
1: /cgn2_6/prodata/1/1aa/5A_COMB pep: *
2: /cgn2_6/prodata/1/1aa/5E_COMB pep: *
3: /cgn2_6/prodata/1/1aa/5F_COMB pep: *
4: /cgn2_6/prodata/1/1aa/5G_COMB pep: *
5: /cgn2_6/prodata/1/1aa/PTCT5_COMB pep: *
6: /cgn2_6/prodata/1/1aa/backfiles1 pep: *
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	length	DB	ID	Description
1	79.5	8.0	202	1	US-07-807-022a-1	Sequence 1, Appl1	
2	78	7.9	1095	4	US-09-107-532a-3855	Sequence 3855, Ap	
3	77.5	7.8	468	4	US-09-107-532a-5304	Sequence 5304, Ap	
4	77.5	7.8	823	4	US-09-107-532a-5667	Sequence 5667, Ap	
5	77	7.8	341	4	US-09-563-269-19	Sequence 19, Appl	
6	76.5	7.7	216	3	US-08-824-692-24	Sequence 24, Appl	
7	75.5	7.6	347	4	US-09-328-352-6764	Sequence 6764, Ap	
8	75	7.6	500	4	US-09-071-035-3396	Sequence 396, Appl	
9	75	7.6	1074	4	US-09-071-035-3358	Sequence 358, Appl	
10	75	7.6	1074	4	US-09-071-035-3394	Sequence 394, Appl	
11	74.5	7.5	334	4	US-09-198-458-1120	Sequence 1120, Ap	
12	74	7.5	771	4	US-09-634-238-229	Sequence 229, Appl	
13	74	7.5	771	4	US-09-634-238-262	Sequence 262, Appl	
14	73.5	7.4	144	4	US-09-252-918a-23258	Sequence 23258, Ap	
15	73	7.4	325	3	US-08-961-083-80	Sequence 80, Appl	
16	73	7.4	325	4	US-09-566-784-80	Sequence 80, Appl	
17	72.5	7.3	401	4	US-09-232-918a-17272	Sequence 17272, A	
18	72.5	7.3	565	3	US-08-961-083-218	Sequence 218, Appl	
19	72.5	7.3	565	4	US-09-536-784-218	Sequence 218, Appl	
20	72.5	7.3	901	2	US-08-884-681-5	Sequence 5, Appl1	
21	72	7.3	901	3	US-09-258-643-5	Sequence 5, Appl1	
22	72	7.3	1244	5	PCT-US93-10500-2	Sequence 2, Appl1	
23	71.5	7.2	397	1	US-08-290-448a-76	Sequence 76, Appl	
24	71.5	7.2	397	1	US-08-290-448a-76	Sequence 76, Appl	
25	71.5	7.2	397	1	US-08-175-069a-76	Sequence 76, Appl	
26	71.5	7.2	397	4	US-08-461-939b-76	Sequence 76, Appl	
27	71.5	7.2	397	4	US-08-464-000-76	Sequence 76, Appl	

## ALIGNMENTS

28	71.5	7.2	805	4	US-09-134-001C-4821	Sequence 4821, App1
29	71	7.2	620	4	US-08-637-670-40	Sequence 40, App1
30	71	7.2	648	4	US-09-199-637A-221	Sequence 221, App1
31	71	7.2	972	3	US-08-1335-844A-24	Sequence 24, App1
32	71	7.2	972	4	US-09-129-366-24	Sequence 24, App1
33	70.5	7.1	240	3	US-08-824-692-23	Sequence 23, App1
34	70.5	7.1	816	4	US-09-107-5324-3862	Sequence 3862, App1
35	70	7.1	246	3	US-09-113-7050A-34	Sequence 34, App1
36	70	7.1	471	4	US-09-134-001C-4125	Sequence 4125, App1
37	70	7.1	603	4	US-09-508-264A-1	Sequence 1, App1
38	70	7.1	845	4	US-08-651-567B-37	Sequence 37, App1
39	70	7.1	851	4	US-09-071-035-326	Sequence 326, App1
40	70	7.1	851	4	US-09-071-035-330	Sequence 330, App1
41	70	7.1	851	4	US-09-071-035-334	Sequence 334, App1
42	70	7.1	862	4	US-09-206-551-15	Sequence 15, App1
43	70	7.1	890	1	US-08-472-934-8	Sequence 8, App1
44	70	7.1	890	1	US-08-323-460A-8	Sequence 8, App1
45	70	7.1	890	2	US-08-461-146C-8	Sequence 8, App1

```

1
US-07-807-022A-1
: Sequence 1, Application US/07807022A
: Patent No. 5321010
:
: GENERAL INFORMATION:
: APPLICANT: Connolly, Thomas M.
: APPLICANT: Neepser, Michael
: APPLICANT: Waxman, Lloyd
: TITLE OF INVENTION: Protein for Inhibiting Adhesion of
: TITLE OF INVENTION: Platelets to Collagen
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Merck & Co., Inc.
: STREET: P.O. Box 2000
: CITY: Rahway
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 08840
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/807,022A
: FILING DATE: 19911210
: CLASSIFICATION: 514
:
: ATTORNEY/AGENT INFORMATION:
: NAME: Parr, Richard S.
: REGISTRATION NUMBER: 32,586
: REFERENCE/DOCKET NUMBER: 18143
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (908) 594-4958
: TELEFAX: 908-594-4720
: TELEEX: 138825
:
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 202 amino acids
: TYPE: amino acid
: STRANDEDNESS: unknown
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
:
: US-07-807-022A-1.

```

Db 33 RDIKKT--YKSKYFVFLIRRTYWD---GRSLCRYGTVLK--RDKANHRI---EQINGMS 82  
Qy 74 LKESHEITVWKAYDYT-----TENGIKYETQTRQTFEDVFEVSDKMDYIFVPE 127  
Db 83 YDESEYETKTKVYITFRIGSGDRNHGVSLOQYVNHGIEYKMIYDDOCALIKATYKD 142  
Qy 128 RGSDF---EGDYELWVSEDKIDKIPDCKFTMAVFAOQOEXTVRNVYTD---SSCFAP 179  
Db 143 NRNPONLKNACEMWATANDANSVSIACEVY---ORCNPNNSVDFFPYVTCKWPP 199  
Qy 180 AQ 181  
Db 200 VE 201  
RESULT 2  
US-09-107-532A-3855  
Sequence 3855, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Arinello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 3855:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1095 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (B) LOCATION 1...1095  
SEQUENCE DESCRIPTION: SEQ ID NO: 3855:  
US-09-107-532A-3855  
Query Match 7.98; Score 78; DB 4; Length 1095;  
Best Local Similarity 23.08; Pred. No. 12;  
Matches 49; Conservative 34; Mismatches 64; Indels 66; Gaps 13;  
Qy 15 DANKSLQD--DUNKRYVLAOTOT---TDGVWGEFTCVSVTAERIGKKLNAT--- 64  
Db 819 DAIKAQDEKDEKAKPVYIAETSAITLANKETGTWIKIOH---KLTAEQV---LKKITVLF 871

Qy 65 --LYKNKH-----LTDKESHEITVWKAYDYTTENGIKYETQ-----GTRQTFE 108  
Db 872 NYVENKNAPFAGGKPAKADVSLNNQATV-----SCYEHVSIQTFAHLNMSGOTTHG 927  
Qy 109 DVF-VFSDYK-NCVYIFPKERGS-----DEGDYELW-----VSE 141  
Db 928 DVMDFDVSITTHVLGSKKEAFETIYALLPDETNRKIMWSKIEYEVDNKEFTKYVLA 987  
Qy 142 DKID--KIPDCKFTMAVFAOQOEXTVRNVYTD 172  
Db 988 KKVDTGKTPDGKTFPAEINVDKDTINGKNE 1020  
RESULT 3  
US-09-107-532A-5304  
Sequence 5304, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Arinello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 5304:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 468 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (B) LOCATION 1...468  
SEQUENCE DESCRIPTION: SEQ ID NO: 5304:  
US-09-107-532A-5304  
Query Match 7.88; Score 77.5; DB 4; Length 468;  
Best Local Similarity 24.18; Pred. No. 4.1;  
Matches 32; Conservative 16; Mismatches 50; Indels 35; Gaps 4;  
Qy 12 SYDANKSLQDQDUNKRYVLAOTQTPDGVWGEFTCVSVTAERIGKKLNA-TILYKNKH 70  
Db 12 NYPLTRPDKOKLKRPIYRSLQD-----LEQDILKNLQKNQOLQPSQRE 56

[illegible][illegible]

REFERENCE/DOCKET NUMBER: 130001.404  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 216 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-824-692-24

Query Match  
Best Local Similarity 19.5%; Pred. No. 1.8;  
Matches 36; Conservative 31; Mismatches 69; Indels 49; Gaps 9;

QY 19 SLQDDOKKRYLAQATQ-----TTGCGWGEER-TCVSTAEK---IGK 57  
Db 47 SSAMEPREYHFGAVFVCSGYKIEGDEEMHCSDFGFGKPKCWEISCKSPDYING 106  
QY 58 KLNATILYKKNKHLTDLKESHETITWKAYDTTENGIRKYEOTGR--TQTFEDVFFVPSD 115  
Db 107 SPIQKIIYKEN-----ERFOYKCNMGIEYSEKDAVCTESGMRPLPSC 151  
QY 116 YKNCQVITFVKERGSDEGDIY-LWVSEDKIDKIPDCKFTMAYPAQOQEKTVRNVYDSS 174  
Db 152 EKSCDNPYIP-----NGDYSPILKIKRTGDEITYGCR--NGEYFATRGNTAK--CSTG 201  
QY 175 CKRAP 179  
Db 202 WIRAP 206

RESULT 7  
US-09-328-352-6764  
Sequence 6764, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: GCG99-03PA  
CURRENT APPLICATION NUMBER: US/09/328.352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 6764  
LENGTH: 347  
TYPE: PRP  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-6764

Query Match  
Best Local Similarity 22.2%; Pred. No. 4.5;  
Matches 24; Conservative 19; Mismatches 42; Indels 23; Gaps 4;

QY 6 NEAKLSYODAMKSLQ--DQNKRYLAQATQTTDGWGEETFCVSTAEKIGKKLNAT 63  
Db 234 NMSHLTAHQEEMKKIIQWMDKTYKINDPATHA-----DAVKIMARSGVDPRQYE 264  
QY 64 ILKKNKHLTDLKESHETITWKAYD-----YTTENGIRKYEIQ 100  
Db 285 LMSVSGTHLDINANKKVFYSQGFDSYSSYHNKFNENBNGI-YKTE 331

RESULT 8  
US-09-071-035-396  
Sequence 396, Application US/09071035  
Patent No. 6448043  
GENERAL INFORMATION:  
APPLICANT: Gil H. Choi  
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
NUMBER OF SEQUENCES: 496  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071.035  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
ATTORNEY/AGENT INFORMATION:  
NAME: A. Anders Brookes  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB369P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 396:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 500 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-071-035-396

Query Match  
Best Local Similarity 28.2%; Pred. No. 8.5;  
Matches 22; Conservative 13; Mismatches 29; Indels 14; Gaps 3;

QY 11 GSYODAMKS--LQDDOKRYR---LAQATQTTDGWGEETFCVSTAEKIGKKLNAT 64  
Db 400 GINKELIKSKKIEHVDKKEFTYVLAEKYDTGKYPGKIFTTETINYEKDN----- 452  
QY 65 LYKKNKHLTDLKESHETIT 82  
Db 453 -VNGKHNEDLKESQTLF 469

RESULT 9  
US-09-071-035-358  
Sequence 358, Application US/09071035  
Patent No. 6448043  
GENERAL INFORMATION:  
APPLICANT: Gil H. Choi  
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
NUMBER OF SEQUENCES: 496  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071.035  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

FILED DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: A. Anders Brookes  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB369P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 358:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1074 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-071-035-358

Query Match 7.6% Score 75; DB 4; Length 1074;  
Best Local Similarity 28.2% Pred. No. 25;  
Matches 22; Conservative 13; Mismatches 29; Indels 14; Gaps 3;

OY 11 GSYDAMKS--LQDDQNKRY--LAQATOTDGVGGEFTCVSVTAERIGKKLNATI 64  
DB 940 GTNKEIWKSGRIEHEVNDKEFTKTVLEKVDGTGKYPGEKFTFTEINEXKDN----- 992

OY 65 LYKNKHLTDLKESHETIT 82  
DB 993 -VNGKHNDLKEKSQTILT 1009

RESULT 10  
US-09-071-035-394

Sequence 394, Application US/09071035  
Patent No. 6448043  
GENERAL INFORMATION:  
APPLICANT: Gil H. Choi  
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
NUMBER OF SEQUENCES: 496  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 Inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071.035  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: A. Anders Brookes  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB369P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 394:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1074 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-071-035-394

Query Match 7.6% Score 75; DB 4; Length 1074;  
Best Local Similarity 28.2% Pred. No. 25;  
Matches 22; Conservative 13; Mismatches 29; Indels 14; Gaps 3;

OY 11 GSYDAMKS--LQDDQNKRY--LAQATOTDGVGGEFTCVSVTAERIGKKLNATI 64  
DB 940 GTNKEIWKSGRIEHEVNDKEFTKTVLEKVDGTGKYPGEKFTFTEINEXKDN----- 992

OY 65 LYKNKHLTDLKESHETIT 82  
DB 993 -VNGKHNDLKEKSQTILT 1009

RESULT 11  
US-09-198-452A-1120

Sequence 1120, Application US/09198452A  
Patent No. 6559294  
GENERAL INFORMATION:  
APPLICANT: Griffiths, R.  
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragment thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection  
TITLE OF INVENTION: and treatment of infection  
FILE REFERENCE: 9710-003-999  
CURRENT FILING DATE: 1998-11-24  
NUMBER OF SEQ ID NOS: 6849  
SEQ ID NO 1120  
LENGTH: 334  
TYPE: PRT  
ORGANISM: Chlamydia pneumoniae  
US-09-198-452A-1120

Query Match 7.5% Score 74.5; DB 4; Length 334;  
Best Local Similarity 28.4% Pred. No. 5.4;  
Matches 29; Conservative 17; Mismatches 37; Indels 19; Gaps 6;

OY 5 ANEAKLSYDAMKSLQDDQNKRY--LAQATOTDGVGGEFTCVSVTAERIGKKKL 60  
DB 97 ATRVCGLA---AMNAND---RFDRLANVKSITD--LGAIVCALGMLSEBQAKLY 148

OY 61 NATIYKNKHLTDLKESHETITVKAY-----DYTENGCI 95  
DB 149 DAGLYAVYVNHNDSSPEFYETIITRSYEDRLNTLDVNVKSGI 190

RESULT 12  
US-09-634-238-229

Sequence 229, Application US/09634238  
Patent No. 6544772  
GENERAL INFORMATION:  
APPLICANT: Glenn, Matthew  
APPLICANT: Havukkala, Ilkka J.  
APPLICANT: Bloksberg, Leonard, N.  
APPLICANT: Lubbers, Mark W.  
APPLICANT: Dekker, James  
APPLICANT: Christensen, Anna C.  
APPLICANT: Holland, Ross  
APPLICANT: O'Toole, Paul W.  
APPLICANT: Reid, Julian R.  
TITLE OF INVENTION: Polynucleotides, materials incorporating  
TITLE OF INVENTION: them and methods for using them.  
FILE REFERENCE: 11000.104301  
CURRENT APPLICATION NUMBER: US/09/634.238  
CURRENT FILING DATE: 2000-08-08  
NUMBER OF SEQ ID NOS: 422  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 229  
LENGTH: 771  
TYPE: PRT  
ORGANISM: Lactobacillus rhamnosus  
US-09-634-238-229



Db	216	HKMNGDRYNQFAGLRNLVYQICHGKKLFM---- <td>271</td>	271
Oy	151	CKFTMAYFAOQOEKTVRNVYTDSSC	175
Db	272	MNAKMKYFASO----LNQFYKDHRG	292

Search completed: August 1, 2003, 12:17:22  
Job time : 15.8667 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2003, 12:15:04 ; Search time 56.5067 Seconds  
(without alignments)  
511.236 Million cell updates/sec

Title: US-10-087-195-8  
Perfect score: 989  
Sequence: 1 NPTWANEAKLSYODAMKSL.....EKTVRNVTYDSSCKPAPAQN 182

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues  
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- A.GeneSeq.19Jun03:\*
- 1: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:\*
  - 2: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:\*
  - 3: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:\*
  - 4: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:\*
  - 5: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:\*
  - 6: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:\*
  - 7: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:\*
  - 8: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:\*
  - 9: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:\*
  - 10: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:\*
  - 11: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:\*
  - 12: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:\*
  - 13: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:\*
  - 14: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:\*
  - 15: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:\*
  - 16: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:\*
  - 17: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:\*
  - 18: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:\*
  - 19: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:\*
  - 20: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:\*
  - 21: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:\*
  - 22: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:\*
  - 23: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:\*
  - 24: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	989	100.0	182	AA174290	Histacalin protein
2	989	100.0	182	AA173262	Histacalin protein
3	989	100.0	200	AA174448	Tick vasoactive am
4	989	100.0	200	AA18080	Histamine binding
5	331.5	33.5	171	AA174289	Histacalin protein
6	331.5	33.5	171	AA173261	Histacalin protein
7	331.5	33.5	190	AA174447	Tick vasoactive am
8	331.5	33.5	190	AA18079	Histamine binding
9	267.5	27.0	172	AA174288	Histacalin protein

10	267.5	27.0	172	AA173260	Histacalin protein
11	267.5	27.0	190	AA174446	Tick vasoactive am
12	267.5	27.0	190	AA18078	Histamine binding
13	267	27.0	203	AA18086	Histamine binding
14	263	26.6	203	AA18085	Histamine binding
15	251.5	25.4	209	AA174449	Tick vasoactive am
16	251.5	25.4	209	AA18081	Histamine binding
17	243	24.6	198	AA18084	Histamine binding
18	123.5	12.5	285	AA18087	Histamine binding
19	119.5	12.1	284	AA18088	Histamine binding
20	109.5	11.1	161	ABP04842	Human ORP protein
21	101.5	10.3	207	AA18083	Histamine binding
22	92.5	9.4	321	AA18089	Histamine binding
23	87	8.8	483	AB184292	Listeria monocylog
24	85	8.6	514	AB184484	Theobroma cacao as
25	85	8.6	1959	AB184411	Listeria monocylog
26	83.5	8.4	778	AA181217	Enterococcus faeca
27	82.5	8.3	1802	AA181217	Plasmodium falcipa
28	82	8.3	392	ABP30105	Streptococcus poly
29	80.5	8.1	721	AB184955	Listeria monocylog
30	79.5	8.0	202	AA183467	Sequence of prote
31	79.5	8.0	2681	AB183025	Pathogen specific
32	79	8.0	221	AA197764	I. scapularis Salp
33	79	8.0	375	AB184335	Listeria monocylog
34	79	8.0	396	AA1825660	Ragweed allergen 1
35	79	8.0	431	AA1827933	Human secreted pro
36	79	8.0	541	AA1827995	Human secreted pro
37	79	8.0	542	AA1827996	Human secreted pro
38	79	8.0	1491	AA1823209	Human secreted pro
39	79	8.0	1605	AA1823182	Novel human secret
40	79	8.0	1607	AA1823210	Human mitogen-acti
41	79	8.0	1830	AA1844504	Urea amide lyase.
42	78.5	7.9	429	AB173831	Candida albicans e
43	77.5	7.8	423	AB1847375	Listeria monocylog
44	77	7.8	341	AA1836492	Bacillus thuringie
45	77	7.8	745	AB175664	Human secretory po

ALIGNMENTS

RESULT 1	AA174290	standard; protein; 182 AA.
ID	AA174290	
AC	AA174290:	
DT	20-JUN-2001	(first entry)
DE	Histacalin protein MS-HBPL.	
KW	Histacalin; MS-HBPL; conjunctivitis.	
OS	Unidentified.	
PN	WO200115719-A2.	
PD	08-MAR-2001.	
PF	24-AUG-2000; 2000WO-GB03282.	
PR	01-SEP-1999; 99GB-0020674.	
PA	(EVOL-) EVOLUTEC LTD.	
XX	Nuttall PA, Paesen GC;	
XX	WPI: 2001-257675/26.	
PT	Use of histacalin proteins for treating or preventing non-infective	
PT	conjunctivitis, or for manufacturing a medicament for treating or	
PT	preventing conjunctivitis, e.g. seasonal or perennial allergic	
PT	conjunctivitis	

XX Claim 4; Page 5-6; 19pp; English.  
XX  
CC The present invention relates to the use of a histacalin protein  
CC for treating or preventing conjunctivitis. The present sequence  
CC is the histacalin protein PS-HBPI. The invention is particularly useful  
CC in the treatment of allergic or seasonal conjunctivitis.  
XX  
SQ Sequence 182 AA;

Query Match 100.0%; Score 989; DB 22; Length 182;  
Best Local Similarity 100.0%; Pred. No. 3,1e-94;  
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPTWANEAKLGSYODANKSLQODONKRYRYLAQATQTTDVGWGEFFTCVSTAERIGKKL 60  
DB 1 NPTWANEAKLGSYODANKSLQODONKRYRYLAQATQTTDVGWGEFFTCVSTAERIGKKL 60  
QY 61 NATLYNKNHLLTDKESHEITTWKAYDYTTENGIRKETGTRQTFEDVFEVSDYKNC 120  
DB 61 NATLYNKNHLLTDKESHEITTWKAYDYTTENGIRKETGTRQTFEDVFEVSDYKNC 120  
QY 121 VIFVPERGSDGDEGYELWVSEDKIDKIPDCCKFTMAVFAOQOEKTVANNVYTDSSCKRPA 180  
DB 121 VIFVPERGSDGDEGYELWVSEDKIDKIPDCCKFTMAVFAOQOEKTVANNVYTDSSCKRPA 180  
QY 181 QN 182  
DB 181 QN 182

## RESULT 2

ID AAB73262 standard; protein; 182 AA.

AC AAB73262;

DT 15-MAY-2001 (first entry)

DE Histacalin protein MS-HBPI.

XX MS-HBPI; histacalin protein; antinflammatory; antiallergic;

KW ophthalmological; allergic rhinitis.

XX Unidentified.

OS WO200116164-A2.

PN 08-MAR-2001.

PD 24-AUG-2000; 2000WO-GB03287.

PE 01-SEP-1999; 99GB-0020673.

PR (EVOL-) EVOLUTEC LTD.

PA Nuttall PA, Paesen GC;

XX Nuttall PA, Paesen GC;

PI WPI; 2001-218521/22.

DR WPI; 2001-218521/22.

XX Use of histacalin proteins for treating or preventing allergic

PT rhinitis, or for manufacturing a medicament for treating or preventing

PT allergic rhinitis, e.g. seasonal or perennial allergic rhinitis

XX Disclosure; Pages 4-6; 19pp; English.

XX The present invention relates to a method for treating or preventing

XX allergic rhinitis. The method involves employing a blood-feeding

XX ectoparasite-derived (e.g. tick-derived) histacalin protein. The present

XX sequence is one such histacalin protein. The histacalin protein, is

XX useful for treating or preventing allergic rhinitis, both seasonal and

XX perennial allergic conjunctivitis.

SQ Sequence 182 AA;

Query Match 100.0%; Score 989; DB 22; Length 182;  
Best Local Similarity 100.0%; Pred. No. 3,1e-94;  
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPTWANEAKLGSYODANKSLQODONKRYRYLAQATQTTDVGWGEFFTCVSTAERIGKKL 60  
DB 1 NPTWANEAKLGSYODANKSLQODONKRYRYLAQATQTTDVGWGEFFTCVSTAERIGKKL 60  
QY 61 NATLYNKNHLLTDKESHEITTWKAYDYTTENGIRKETGTRQTFEDVFEVSDYKNC 120  
DB 61 NATLYNKNHLLTDKESHEITTWKAYDYTTENGIRKETGTRQTFEDVFEVSDYKNC 120  
QY 121 VIFVPERGSDGDEGYELWVSEDKIDKIPDCCKFTMAVFAOQOEKTVANNVYTDSSCKRPA 180  
DB 121 VIFVPERGSDGDEGYELWVSEDKIDKIPDCCKFTMAVFAOQOEKTVANNVYTDSSCKRPA 180  
QY 181 QN 182  
DB 181 QN 182

## RESULT 3

ID AAB7448 standard; protein; 200 AA.

AC AAB7448;

DT 08-JUN-1998 (first entry)

DE Tick vasocactive amine binding protein 1 MS-HBPI.

XX Male-specific vasocactive amine binding protein 1; MS-HBPI;

KW histamine; serotonin; assay; antihistamine; anti-inflammatory;

KW insect bite; snake bite; scorpion bite; dermatitis; vaccine;

XX transgenic animal; tick.

XX Rhipicephalus appendiculatus.

OS Rhipicephalus appendiculatus.

XX Key Location/Qualifiers

FT Peptide 1..18

FT Modified-site 79..81

FT /label= sig-peptide

FT /note= "Asn is N-glycosylated"

FT WO9744451-A2.

PN 27-NOV-1997.

PD 19-MAY-1997; 97WO-GB01372.

PE 18-APR-1997; 97GB-0007844.

PR 18-MAY-1996; 96GB-0010484.

PA (OXFO-) OXFORD VACS LTD.

XX Nuttall PA, Paesen GC;

PI WPI; 1998-018506/02.

DR N-PSDB; AAV00229.

XX New vasocactive amine binding proteins and related nucleic acid,

PT vectors - transformed cells and transgenic animals, used for

PT assaying or removing histamine and as antihistamine or

PT anti-inflammatory agents

XX Example 2; Fig 3; 44pp; English.

XX This protein comprises tick Rhipicephalus appendiculatus (Ra) novel

XX male-specific histamine binding protein 1 (MS-HBPI). Its amino

XX acid sequence was deduced from a cDNA clone (see AAV00227) obtained

XX from a salivary gland cDNA library. Female-specific HBPI and

CC HSP2 (see AAW37446-47) and a related protein, D-RET6 (see AAW37449)  
 CC from Dermacenter reticulatus, were also identified. These novel  
 CC vasoactive amine binding proteins (VABPs) can be expressed in  
 CC host cells using e.g. a baculovirus expression system. They can  
 CC be used: (i) to assay histamine (or other VA such as serotonin) in  
 CC body fluids or cell culture supernatants, e.g. to monitor the  
 CC effect of allergens; (ii) for binding VA, e.g. to remove histamine  
 CC from blood, food, cell cultures etc.; (iii) as an antihistamine or  
 CC anti-inflammatory agents, e.g. for treating insect, snake or  
 CC scorpion bites or dermatitis; or as a carrier for slow release of  
 CC histamine-related compounds; (iv) in vaccines to protect against  
 CC metazoan parasites, especially in animals; (v) as reagents for  
 CC studying inflammation, involvement of VA in ulcer formation or the  
 CC immune response etc. VABPs provide a more sensitive assay for  
 CC histamine than low-affinity antibodies currently used. They may  
 CC also be more effective and safer than conventional antihistamines.  
 CC  
 CC Sequence 200 AA:

Query Match 100.0%; Score 989; DB 19; Length 200;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-94;  
 Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPTVANEAKLGSYODAMKSLQODONKRYLLAOTQTDTGVEEFTCVSYAEKIGKKKL 60  
 DB 19 NPTVANEAKLGSYODAMKSLQODONKRYLLAOTQTDTGVEEFTCVSYAEKIGKKKL 78  
 QY 61 NATLIYKKKHLTDKESHETITWKAYDYTTENGKIKYETOGRTOTFEDEVFESDYKNC 120  
 DB 79 NATLIYKKKHLTDKESHETITWKAYDYTTENGKIKYETOGRTOTFEDEVFESDYKNC 138  
 QY 121 VIFVPERKSGDEGYELWVSEDKIDIPDCKFTMAFYAQOQEKTVRVNYTDSCKPAPA 180  
 DB 139 VIFVPERKSGDEGYELWVSEDKIDIPDCKFTMAFYAQOQEKTVRVNYTDSCKPAPA 198  
 QY 181 QN 182  
 DB 199 QN 200

RESULT 4  
 AAY18080  
 ID AAY18080 standard; Protein: 200 AA.

AC AAY18080;  
 XX  
 DT 06-AUG-1999 (first entry)  
 XX  
 DE Histamine binding protein MS-HBP1.

XX Histamine binding protein; serotonin binding compound; inflammation;  
 KW gastric acid secretion; allergy; type I hypersensitivity reaction;  
 KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
 KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
 KW respiratory disease; coronary heart disease; cellular growth regulator;  
 KW tissue repair; blood-sucking ectoparasite; therapy.

OS Rhipicephalus appendiculatus.

XX W09927104-A1.  
 XX 03-JUN-1999.  
 PD  
 XX 26-NOV-1998; 98WO-GB03530.  
 PF  
 XX 26-JUN-1998; 98GB-0013917.  
 PR 26-NOV-1997; 97GB-0025046.  
 XX  
 XX (OXFORD) OXFORD VACS LTD.  
 PA  
 XX Nuttall PA, Paesen GC;  
 PI  
 XX WPI: 1999-357841/30.  
 DR

DR N-PSDB: AAX76966.  
 XX  
 XX Histamine and serotonin binding compounds useful for the treatment  
 PT of allergies  
 PS  
 XX Claim 1: Fig 3; 84pp; English.

This sequence is an example of a histamine or serotonin binding  
 compound (A), of the invention. The compounds are useful for regulating  
 the action of histamine and serotonin (in e.g. inflammation and gastric  
 acid secretion), the detection, quantification and removal of histamine  
 or serotonin (in animals, plants, cell cultures, food materials), or  
 humans) and in the treatment of various diseases and allergies  
 (e.g. type I hypersensitivity reactions, urticaria, asthma, allergic  
 rhinitis (hay fever), atopic dermatitis, insect bites and food and drug  
 allergies, abnormal blood pressure, migraine, psychological disorders,  
 respiratory disease, and coronary heart disease). Histamine may also be  
 used to regulate cellular growth and tissue repair. The molecules may  
 also be used as components of vaccines directed against blood-sucking  
 ectoparasites.

Query Match 100.0%; Score 989; DB 20; Length 200;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-94;  
 Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPTVANEAKLGSYODAMKSLQODONKRYLLAOTQTDTGVEEFTCVSYAEKIGKKKL 60  
 DB 19 NPTVANEAKLGSYODAMKSLQODONKRYLLAOTQTDTGVEEFTCVSYAEKIGKKKL 78  
 QY 61 NATLIYKKKHLTDKESHETITWKAYDYTTENGKIKYETOGRTOTFEDEVFESDYKNC 120  
 DB 79 NATLIYKKKHLTDKESHETITWKAYDYTTENGKIKYETOGRTOTFEDEVFESDYKNC 138  
 QY 121 VIFVPERKSGDEGYELWVSEDKIDIPDCKFTMAFYAQOQEKTVRVNYTDSCKPAPA 180  
 DB 139 VIFVPERKSGDEGYELWVSEDKIDIPDCKFTMAFYAQOQEKTVRVNYTDSCKPAPA 198  
 QY 181 QN 182  
 DB 199 QN 200

RESULT 5  
 AAB74289  
 ID AAB74289 standard; Protein: 171 AA.

AC AAB74289;  
 XX  
 DT 20-JUN-2001 (first entry)  
 XX  
 DE Histacalin protein FS-HBP2.

XX Histacalin; FS-HBP1; conjunctivitis.

OS Undentified.

XX W0200115719-A2.  
 XX 08-MAR-2001.  
 PD  
 XX 24-AUG-2000; 2000MO-GB03282.  
 PF  
 XX 01-SEP-1999; 99GB-0020674.  
 PR  
 XX (EVOL) EVOLUTEC LTD.  
 PA  
 XX Nuttall PA, Paesen GC;  
 PI  
 XX WPI: 2001-257675/26.  
 DR  
 XX Use of histacalin proteins for treating or preventing non-infective  
 PT

PT conjunctivitis, or for manufacturing a medicament for treating or  
 PT preventing conjunctivitis, e.g. seasonal or perennial allergic  
 PT conjunctivitis  
 PS Claim 4; Page 5-6; 19pp; English.  
 XX  
 CC The present invention relates to the use of a histacalin protein  
 CC for treating or preventing conjunctivitis. The present sequence  
 CC is the histacalin protein FS-HBP1. The invention is particularly useful  
 CC in the treatment of allergic or seasonal conjunctivitis.  
 XX  
 SQ Sequence 171 AA;  
 Query Match 33.5%; Score 331.5; DB 22; Length 171;  
 Best Local Similarity 42.0%; Pred. No. 3.6e-26;  
 Matches 74; Conservative 29; Mismatches 62; Indels 11; Gaps 7;  
 QY 2 PTMANEKLGSYQAMKSLQODQKRYIAQATQTGTGWCGEETCYVTAERKIG--KKK 59  
 DB 3 PDMADEANGAHQDAMKSLKADVENYVMKATYKNDPVGNDFTCVGMANDVNEDEKS 62  
 QY 60 LNAITLYKNKHLDTLKSHEITVWKAYDYTTENGIRYETQGTQTFEDEVFVPSDYKNC 119  
 DB 63 IQAEFLPMNADTMQATEKVTAVKMGYRENAFREIED--GQVFTDYIASD-DNC 119  
 QY 120 DVIVPFRKSGDESDYELMWSEDKIDKIPDCKFTMAVFAQOQEKTVRNYTSSC 175  
 DB 120 DVIVPQTDGNEEG-YELMTTD--YDNIIPANCLNKFNEYAVGRE--TRDVFET-SAC 169  
 RESULT 6  
 AAB73261  
 ID AAB73261 standard; protein; 171 AA.  
 AC AAB73261;  
 XX  
 DT 15-MAY-2001 (first entry)  
 XX  
 DE Histacalin protein FS-HBP2.  
 XX  
 KW FS-HBP2; histacalin protein; antiinflammatory; antiallergic;  
 KW ophthalmological; allergic rhinitis.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200116164-A2.  
 PD 08-MAR-2001.  
 XX  
 PF 24-AUG-2000; 2000WO-GB03287.  
 XX  
 PR 01-SEP-1999; 99GB-0020673.  
 XX  
 PA (EVOL-) EVOLUTEC LTD.  
 XX  
 PI Nuttall PA, Paesen GC;  
 DR WPI; 2001-218521/22.  
 XX  
 PT Use of histacalin proteins for treating or preventing allergic  
 PT rhinitis, or for manufacturing a medicament for treating or preventing  
 PT allergic rhinitis, e.g. seasonal or perennial allergic rhinitis -  
 XX  
 PS Disclosure; Pages 4-6; 19pp; English.  
 XX  
 CC The present invention relates to a method for treating or preventing  
 CC allergic rhinitis. The method involves employing a blood-feeding  
 CC ectoparasite-derived (e.g. tick-derived) histacalin protein. The present  
 CC sequence is one such histacalin protein. The histacalin protein, is  
 CC useful for treating or preventing allergic rhinitis, both seasonal and  
 CC perennial allergic conjunctivitis.  
 XX  
 SQ Sequence 171 AA;

Query Match 33.5%; Score 331.5; DB 22; Length 171;  
 Best Local Similarity 42.0%; Pred. No. 3.6e-26;  
 Matches 74; Conservative 29; Mismatches 62; Indels 11; Gaps 7;  
 QY 2 PTMANEKLGSYQAMKSLQODQKRYIAQATQTGTGWCGEETCYVTAERKIG--KKK 59  
 DB 3 PDMADEANGAHQDAMKSLKADVENYVMKATYKNDPVGNDFTCVGMANDVNEDEKS 62  
 QY 60 LNAITLYKNKHLDTLKSHEITVWKAYDYTTENGIRYETQGTQTFEDEVFVPSDYKNC 119  
 DB 63 IQAEFLPMNADTMQATEKVTAVKMGYRENAFREIED--GQVFTDYIASD-DNC 119  
 QY 120 DVIVPFRKSGDESDYELMWSEDKIDKIPDCKFTMAVFAQOQEKTVRNYTSSC 175  
 DB 120 DVIVPQTDGNEEG-YELMTTD--YDNIIPANCLNKFNEYAVGRE--TRDVFET-SAC 169  
 RESULT 7  
 AAM37447  
 ID AAM37447 standard; Protein; 190 AA.  
 AC AAM37447;  
 XX  
 DT 08-JUN-1998 (first entry)  
 XX  
 DE Tick vasoactive amine binding protein 2 FS-HBP2.  
 XX  
 KW Female-specific vasoactive amine binding protein 1; FS-HCP1;  
 KW histamine; serotonin; assay; antihistamine; anti-inflammatory;  
 KW insect bite; snake bite; scorpion bite; dermatitis; vaccine;  
 KW transgenic animal; tick.  
 XX  
 OS Rhipicephalus appendiculatus.  
 XX  
 FH Key location/Qualifiers  
 FT Peptide 1..19  
 FT /label= Sig-peptide  
 XX  
 PN WO9744451-A2.  
 PD 27-NOV-1997.  
 XX  
 PF 19-MAY-1997; 97WO-GB01372.  
 XX  
 PR 18-APR-1997; 97GB-0007844.  
 PR 18-MAY-1996; 96GB-0010484.  
 XX  
 PA (OXFO-) OXFORD VACS LTD.  
 XX  
 PI Nuttall PA, Paesen GC;  
 DR WPI; 1998-018506/02.  
 DR N-PSDB; AAV00228.  
 XX  
 PT New vasoactive amine binding proteins and related nucleic acid,  
 PT vectors - transformed cells and transgenic animals, used for  
 PT assaying or removing histamine and as antihistamine or  
 PT anti-inflammatory agents  
 XX  
 PS Example 2; Fig 2; 44pp; English.  
 XX  
 CC This protein comprises tick Rhipicephalus appendiculatus (Ra) novel  
 CC female-specific histamine binding protein 2 (FS-HBP2). Its amino  
 CC acid sequence was deduced from a cDNA clone (see AAV00228) obtained  
 CC from a salivary gland cDNA library. FS-HSP1 (see AAM37446) and  
 CC male-specific HSP1 (see AAM37448) and a related protein, D.RET6 (see  
 CC AAM37449) from Dermacentor reticulatus, were also identified. These  
 CC novel vasoactive amine binding proteins (VABPs) can be expressed  
 CC in host cells using e.g. a baculovirus expression system. They can  
 CC be used: (i) to assay histamine (or other VA such as serotonin) in  
 CC body fluids or cell culture supernatants, e.g. to monitor the  
 CC effect of allergens; (ii) for binding VA, e.g. to remove histamine





```

0y      2  PYNANEKIGSGYODAKRSLC000ONKRYUAAOQTOTTDG0WGEELFCVSATKEKIG--KKK  59
      ||| ||| : ||||| ||| : ||| ||| ||| ||||| ||| : ||| |||
Db      21  PYADEXANGEHODAKRHLOKLEENRYTDILKATRYKNDPAGWGDFTCGYTAQAONLNDEKN  80
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
0y      60  LNTATLYKNKHLTDLKESHETIYMKAYDTTENGKIKETOGTROPEDEYFVPSDKNC  119
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      81  VEMAFEMNNADTVYHTEFEKATPDKMTGYKNENAIYYQTEB--GQVLTDLVAISD--DNC  137
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
0y      120  DVIFVPERKSGDEGDEYELWSEDKIDKIPDCKFTMAFPAQOQEKTYVNNVTDSSCKP  177
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      138  YVYIALGPDGSGAG-YELMAT--DYTDVPASCLCEKNEETAA---GLPVDVYT--SDCLP  189
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 12  
AAV18078  
ID AAV18078 standard; Protein; 190 AA

AC AAY18078;

DT 06-AUG-1999 (first entry)

### DE Histamine binding protein FS-HBP1.

KW Histamine binding protein; serotonin binding compound; inflammation; gastric acid secretion; allergy; type I hypersensitivity reaction; KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy; KW drug allergy; abnormal blood pressure; psychological disorder; vaccine; KW respiratory disease; coronary heart disease; cellular growth regulator; KW tissue repair; blood-sucking ectoparasite; therapy.

*Rhipicephalus appendiculatus*.

PN WO9927104-A1.

PD 03-JUN-1999.

PF 26-NOV-1998; 98WO-GB03530.

PR 26-JUN-1998; 98GB-0013917.

[illegible]

XX  
XX  
XX

PL NULLELL PA, PAGES  
XX

DR WP1; 1999-35/841/30-  
DR N-PSDB: AAX76964

PT Histamine and serotonin binding compounds useful for the treatment  
XX

PT of allergies  
 YV

PS Claim 1; Fig 1; 84pp; English.

This sequence is an example of a histamine or serotonin binding

the action of histamine and serotonin (in e.g. inflammation and gastric acid secretion), the detection, quantification and removal of histamine or serotonin (in animals, plants, cell cultures, food materials, or humans) and in the treatment of various diseases and allergies (e.g. type I hypersensitivity reactions, urticaria, asthma, allergic rhinitis (hay fever), atopic dermatitis, insect bites and food and drug allergies, abnormal blood pressure, migraine, psychological disorders, respiratory disease, and coronary heart disease). Histamine may also be used to regulate cellular growth and tissue repair. The molecules may also be used as components of vaccines directed against blood-sucking ectoparasites.

Sequence 190 AA;

Query Match	27.0%	Score	267.5	DB	20	Length	190
Best Local Similarity	39.3%	Pred. No.	1.8e-19				
Matches	70	Conservative	20	Mismatches	77	Indels	11
						Gaps	7

```

QY 2 PTWANEAKGSGVQDAMKSLQODONKRYLYAQAOTQTDTGVMGBEEFTCVSYAEKIG--KKK 59
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 21 PVMADPEANGEHQDAMKHLQKLVVEENYDIIKATYKNDPVMGNFTCVGTAAQNLNDEKN 80
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 60 LNATILLYKKKHLJTDKESHETITVMKAYDYTTENGKIKYETQGRRTQTFEDVFVFSYKNC 119
  ::||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 81 VEAMFMFMNADTVYOHFFEKATPDDMYGKNNENATYYTED--GOVLTVLAFSD--DNC 137
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 120 DVIETPKERGSDEGDELTWVSEDKIDKIPDCCKFTMAYFAQOQKEITVRNYYTSSCP 177
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 138 YVITVLGPDGCSAG-YELMNT-DYTDVPAWSCLEKENNEYA--GLPVRDYT--SDCLP 189

```

RESULT 13
AA18086
ID AA18086 standard; Protein; 203 AA

AC MAY18086;

DT 06-AUG-1999 (first entry)

### DE Histamine binding protein Ih/Bm-HBP2

KM Histamine binding protein: serotonin binding compound: inflammation;  
 KM gastric acid secretion: allergy: type I hypersensitivity reaction;  
 KM asthma: urticaria: allergic rhinitis: atopic dermatitis: food allergy;  
 KM drug allergy: abnormal blood pressure: psychological disorder: vaccine  
 KM respiratory disease: coronary heart disease: cellular growth regulator  
 KM tissue repair: blood-sucking ectoparasite: therapy.

OS Boophilus microplus

Synthetic.

PN WO9927104-A1.

PD 03-JUN-1999.

AA 26-NOV-1998; 98WO-GB03530.

PR 26-JUN-1998; 98GB-0013917.

XX 20-NOV-1991 0705-002

PA (OXFORD) OXFORD VACS LTD  
XX

PI Nuttall PA, Paesen  
XX

DR WPI; 1999-357841/30.  
DR N-PSDB: AAY76971

XX  
XX  
XX

PT of allergies

PS Claim 14; Fig 8; 84pp; English.

This sequence is an example of a histamine or serotonin binding compound (A), of the invention. cDNA encoding this sequence was isolated from a mixed Bopophilus microplus/Ixodes hexagonus cDNA expression library. The compounds are useful for regulating the action of histamine and serotonin (in e.g. inflammation and gastric acid secretion), the detection, quantification and removal of histamine or serotonin (in animals, plants, cell cultures, food materials, or humans) and in the treatment of various diseases and allergies (e.g. type I hypersensitivity reactions, urticaria, asthma, allergic rhinitis (hay fever), atopic dermatitis, insect bites and food and drug allergies, abnormal blood pressure, migraine, psychological disorders, respiratory disease, and

CC coronary heart disease  
CC growth and tissue

XX	Sequence	203 AA:
SO		

Query Match 27.0%; Score 267; DB 20; Length 203;  
 Best Local Similarity 36.8%; Pred. No. 2.2e-19;  
 Matches 64; Conservative 26; Mismatches 66; Indels 18; Gaps 8;

QY 3 TWNAEALGSGYODAMKSLQODQNKRYLAQATQTTDGVWGEFFCVSYTA----EKIGK 58  
 DB 34 TWHQOT-LGRQODAMKSLQODQNKRYLAQATQTTDGVWGEFFCVSYTA----EKIGK 58  
 QY 59 KLNATILYKNNHLDLKESHEITVWKAYDYTTENGKYEOTGRTQTFEDVEFESYKRN 118  
 DB 93 EYNTT--YKNSQGVWMS-ENVYAVOEGGYSVANIQTTE--NNKRFNDTVYFTDQOT 147  
 QY 119 CDVIFVP-KERGSDGDEYELWVSEDKIDKIPDCKFTMAVFAQOQEKTVRNVT 171  
 DB 148 CDVILYIPYKEDG-----YELWVRSYLVQNTPTCCQIFDLVA--LGRTYNIST 194

RESULT 14  
 AAY18085  
 ID AAY18085 standard; Protein; 203 AA.  
 AC AAY18085;  
 XX  
 DT 06-AUG-1999 (first entry)  
 XX  
 DE Histamine binding protein Ih/Bm-HBP1.  
 XX  
 KW Histamine binding protein; serotonin binding compound; inflammation;  
 KW gastric acid secretion; allergy; type I hypersensitivity reaction;  
 KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
 KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
 KW respiratory disease; coronary heart disease; cellular growth regulator;  
 KW tissue repair; blood-sucking ectoparasite; therapy.  
 XX  
 OS Boophilus microplus.  
 OS Ixodes hexagonus.  
 OS Synthetic.  
 XX  
 PN WO9927104-A1.  
 PD 03-JUN-1999.  
 XX  
 PF 26-NOV-1998; 98WO-GB03530.  
 XX  
 PR 26-JUN-1998; 98GB-0013917.  
 PR 26-NOV-1997; 97GB-0025046.  
 XX  
 PA (OXFO-) OXFORD VACS LTD.  
 XX  
 PI Nuttall PA, Paesen GC;  
 XX  
 DR WPI; 1999-357841/30.  
 DR N-PSDB; AAX76970.  
 XX  
 PT Histamine and serotonin binding compounds useful for the treatment  
 PT of allergies  
 PS  
 PS Claim 13; Fig 7; 84pp; English.  
 XX  
 CC This sequence is an example of a histamine or serotonin binding  
 CC compound (A), of the invention. cDNA encoding this sequence was isolated  
 CC from a mixed Boophilus microplus/Ixodes hexagonus cDNA expression  
 CC library. The compounds are useful for regulating the action of histamine  
 CC and serotonin (in e.g. inflammation and gastric acid secretion), the  
 CC detection, quantification and removal of histamine or serotonin (in  
 CC animals, plants, cell cultures, food materials, or humans) and in the  
 CC treatment of various diseases and allergies (e.g. type I hypersensitivity  
 CC reactions, urticaria, asthma, allergic rhinitis (hay fever), atopic  
 CC dermatitis, insect bites and food and drug allergies, abnormal blood  
 CC pressure, migraine, psychological disorders, respiratory disease, and  
 CC coronary heart disease). Histamine may also be used to regulate cellular  
 CC growth and tissue repair. The molecules may also be used as components of  
 CC vaccines directed against blood-sucking ectoparasites.

XX  
 SQ Sequence 203 AA;  
 Query Match 26.6%; Score 263; DB 20; Length 203;  
 Best Local Similarity 35.1%; Pred. No. 5.8e-19;  
 Matches 65; Conservative 27; Mismatches 57; Indels 36; Gaps 9;

QY 1 NPTWANAKLGSYODAMKSLQODQNKRYLAQATQTTDGVWGEFFCVSYTA----EKIG 56  
 DB 32 NNTW-HSKELKNYODAMKSLQODQNKRYLAQATQTTDGVWGEFFCVSYTA----EKIG 56  
 QY 57 KKLNATILYKNNHLDLKESHEITVWKAYDYTTENGKYEOTGRTQTF 107  
 DB 91 TVEYNTT--YKNSQGVWMS-ENVYAVOEGGYSVANIQTTE--NNKRFNDTVYFTDQOT 136  
 QY 108 EDVIFVPSDYKNCVIFVP-KERGSDGDEYELWVSEDKIDKIPDCKFTMAVFAQOQEKTV 166  
 DB 137 NDVIFVPSDYKNCVIFVP-KERGSDGDEYELWVSEDKIDKIPDCKFTMAVFAQOQEKTV 189

RESULT 15  
 AAW37449  
 ID AAW37449 standard; Protein; 209 AA.  
 AC AAW37449;  
 XX  
 DT 08-JUN-1998 (first entry)  
 XX  
 DE Tick vasoactive amine binding protein D.RET6.  
 XX  
 KW Vasoactive amine binding protein; D.RET6; histamine; serotonin;  
 KW assay; antihistamine; anti-inflammatory; insect bite; snake bite;  
 KW scorpion bite; dermatitis; vaccine; transgenic animal; tick.  
 XX  
 OS Dermacenter reticularis.  
 OS  
 FH Key Location/Qualifiers  
 FT Peptide 1..28  
 FT /label= sig\_peptide  
 XX  
 PN WO9744451-A2.  
 PD 27-NOV-1997.  
 XX  
 PF 19-MAY-1997; 97WO-GB01372.  
 XX  
 PR 18-APR-1997; 97GB-0007844.  
 PR 18-MAY-1996; 96GB-0010484.  
 XX  
 PA (OXFO-) OXFORD VACS LTD.  
 XX  
 PI Nuttall PA, Paesen GC;  
 XX  
 DR WPI; 1998-018506/02.  
 DR N-PSDB; AAV00230.  
 XX  
 PT New vasoactive amine binding proteins and related nucleic acid,  
 PT vectors - transformed cells and transgenic animals, used for  
 PT assaying or removing histamine and as antihistamine or  
 PT anti-inflammatory agents  
 PS  
 PS Example 2; Fig 4; 44pp; English.  
 XX  
 CC This protein comprises tick Dermacenter reticularis (Dr) novel.  
 CC vasoactive amine binding protein (VABP) D.RET6. Its amino acid  
 CC sequence was deduced from a cDNA clone (see AAV00230) obtained  
 CC from a Dr salivary gland cDNA library. 3 Novel VABPs, designated  
 CC FS-HBP1, FS-HBP2 and MS-HBP1 (see AAW37446-48), of the tick  
 CC Rhipicephalus appendiculatus have also been identified. The



VABPs can be expressed in host cells using, e.g., a baculovirus expression system. They can be used: (i) to assay histamine (or other VA such as serotonin) in body fluids or cell culture supernatants, e.g., to monitor the effect of allergens; (ii) for binding VA, e.g., to remove histamine from blood, food, cell cultures etc.; (iii) as an antihistamine or anti-inflammatory agents, e.g., for treating insect, snake or scorpion bites or dermatitis, or as a carrier for slow release of histamine-related compounds; (iv) in vaccines to protect against metazoan parasites especially in animals; (v) as reagents for studying inflammation, involvement of VA in ulcer formation or the immune response etc. VABPs provide a more sensitive assay for histamine than low-affinity antibodies currently used. They may also be more effective and safer than conventional antihistamines.

**50 Sequence 209 AA;**

Query Match	25.4%	Score 251.5	DB 19	Length 209
Best Local Similarity	35.3%	Pred. No. 9, 3e-18		
Matches 61, Conservative	29	Mismatches 66	Indels 17	Gaps 8

```

OY 1 NPMWANEAKGSGYODAMKRLDODOKRKYLAOTOTOT-GWGGEFTCVSTAJARIKIK 59
Db 30 NPLMAHEBLLGKQDAMKSIDGQSVTYVLAATTYENDGSMGSOFCLOY--QEI8RKE 87
OY 60 LNMATI---LYKKHLITDLKESHENITWMAKDY---TTENGKYEYOGTPTOFEEV 112
Db 88 EDTTVASVTFERNAS-SPIKINYNTETKANAPQYKIRIRAIIEQWGGSLNTI--DTLI 144
OY 113 FSDYKMCVDIYFVPEKESDEGDYELWASEDKITDPDCCKPTMAVFAQOQEXT 165
Db 145 FTGELGCDLFFYV---NADQG--CELMVKKSHYKHHPDCTVEVNFCAKDKRT 193

```

Search completed: August 1, 2003, 12:22:49  
Job time : 57.5067 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2003, 12:15:04 ; Search time 49.1829 seconds  
(without alignments)  
897.203 Million cell updates/sec

Title: US-10-087-195-7

Perfect score: 936  
Sequence: 1 NQPDMADEANGAHQDAWKS.....FNEYAVGRETRDVFYSACLE 171

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	282	30.1	209	5	OBWSK7
2	114.5	12.2	311	5	OBWVC2
3	104.5	11.2	1898	5	OBWL35
4	103	11.0	1590	2	O59983
5	97	10.4	1590	2	O55263
6	96.5	10.3	1449	5	O9GS27
7	92.5	9.9	944	3	O9HEA1
8	92	9.8	176	5	O81979
9	91.5	9.8	221	5	O95W25
10	91.5	9.8	386	16	O8XP49
11	91.5	9.8	640	8	O9GEV3
12	91	9.7	306	5	OBWVC3
13	91	9.7	450	16	O9POA9
14	90.5	9.7	1083	5	O813G1
15	89.5	9.6	542	3	O9P378
16	89.5	9.6	845	10	O9AO25

17	88	9.4	606	16	O977Y3	O977Y3 clostridium
18	87	9.3	377	12	O9OB34	O9OB34 human adeno
19	87	9.3	3381	5	O81DK4	O81DK4 plasmodium
20	86.5	9.2	193	12	O91BS3	O91BS3 turkey herp
21	86.5	9.2	254	5	O23871	O23871 dictyostell
22	86.5	9.2	723	16	O8P1H6	O8P1H6 streptococc
23	86.5	9.2	736	16	O9AOE7	O9AOE7 streptococc
24	86.5	9.2	736	16	O8K7Z8	O8K7Z8 streptococc
25	86	9.2	6077	5	O81CS6	O81CS6 plasmodium
26	85.5	9.1	327	16	O8CP44	O8CP44 staphylococ
27	85.5	9.1	573	5	O81DM2	O81DM2 plasmodium
28	85.5	9.1	617	16	O8X1I6	O8X1I6 clostridium
29	85.5	9.1	686	8	O8WGG0	O8WGG0 dipasac ml
30	85.5	9.1	1237	16	O92CC8	O92CC8 listeria in
31	85.5	9.1	1612	16	O8Y591	O8Y591 listeria mo
32	84	9.0	2488	5	O813R5	O813R5 plasmodium
33	83.5	8.9	4494	5	O81512	O81512 plasmodium
34	83	8.9	601	2	O45821	O45821 clostridium
35	83	8.9	874	2	O9S3L0	O9S3L0 clostridium
36	82.5	8.8	220	5	O8WVB8	O8WVB8 ixodes scap
37	82.5	8.8	496	5	O9CSW0	O9CSW0 plasmodium
38	82	8.8	419	10	O9FJH4	O9FJH4 eradiopsis
39	81.5	8.7	374	16	O8CX67	O8CX67 oceanobacill
40	81.5	8.7	897	5	O81D14	O81D14 plasmodium
41	81.5	8.7	1026	2	O8L2H9	O8L2H9 streptococ
42	81.5	8.7	1129	5	O9VPE3	O9VPE3 drosophila
43	81.5	8.7	1956	5	O81IE1	O81IE1 plasmodium
44	81.5	8.7	3597	5	O81LR5	O81LR5 plasmodium
45	81	8.7	1263	11	O9QY22	O9QY22 mus musculu

## ALIGNMENTS

RESULT 1  
ID O8WSK7 PRELIMINARY: PRT; 209 AA.  
AC O8WSK7:  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Serotonin and histamine binding protein.  
OS Dermacentor reticulatus.  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Parasitiformes; Ixodida; Ixodidae; Dermacentor.  
OX NCBI\_TaxID=57047;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Sangamadech S., Paesen G.C., Nuttall P.A.;  
RT "A high affinity serotonin- and histamine-binding 11pocalin secreted by blood-feeding ticks.";  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF217101; AAL56644.1; -  
DR InterPro; IPR002970; His\_binding.  
DR Pfam; PF02098; His\_binding; 1.  
SQ SEQUENCE 209 AA; 23889 MW; E5EB8BE710859591 CRC64;

Query Match 30.1%; Score 282; DB 5; Length 209;  
Best Local Similarity 37.5%; Pred. No. 5.2e-17;  
Matches 66; Conservative 29; Mismatches 69; Indels 12; Gaps 8;  
Db 3 PMADEANGAHQDAWKS LKADVENYVVKATYKNDP-VWGNDPFCVGMANDVDEDK 61  
1 PMADEANGAHQDAWKS LKADVENYVVKATYKNDP-VWGNDPFCVGMANDVDEDK 61  
62 SIQAEFLPMNNDJTNQF--ATEKYAVVMGY-NRENAFRYETEDGQVFTVIAYSD-D 117  
91 TYSVETP-RNASSPIKYNTVETKAVQYGYKRNRIAEVQVGGLITDTLPTDDE 149  
118 NCDVIVYRGTDNNEGYELMT--TDYDNTPANCLKNFXY-AVGRTRVYFYSACL 170  
150 LCDVYVPMAD--GCELMVKKSHKVPDYCTFVFNVCCKRRTYDIFNEECV 202

```

RESULT 2
Q8MWC2 PRELIMINARY; PRT; 311 AA.
AC 08MWC2:
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Putative secreted histamine binding protein.
DE Ixodes scapularis (Black-legged tick) (Deer tick).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodidae; Ixodidae; Ixodes.
OX NCBI_TaxID=6945;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Rhode Island; TISSUE-Salivary gland;
RA Valenzuela J.G., Francischetti I.M., Pham V.M., Garfield M.,
RA Mather T.N., Ribeiro J.M.C.;
RT "Exploring the Salivome of the Tick Vector of Lyme Disease, Ixodes
RT scapularis."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF483718; AAM93640.1;
SQ SEQUENCE 311 AA; 36581 MW; D7E529FEFEACFA9A CRC64;

Query Match 12.2%; Score 114.5; DB 5; Length 311;
Best Local Similarity 22.8%; Pred. No. 0.041;
Matches 42; Conservative 26; Mismatches 63; Indels 53; Gaps 7;

QY 1 NOPDMADEAANGAHQADAMKSLKADVENYVYVWKATYKNDPYWG-----NDFTCGVAMND 55
DB 31 NNPD-LNRKDLGAMQADAMRIKFTANHSYLIYS-----GNGTREHYEDVRCLOVSSD 84
QY 56 VNDEKSIQAEFLFNANNADTNMOPATEKVTAVKMYGYNRENA----- 97
DB 85 LNTYLKASANYTSKWNRTSKRMNSTQYQAAKQYSIEIMHLAGOPREVTSPNGTCY 144
QY 98 ---FRYTEEDG-----QVFTDVIAYSNDNDVITYPTDNGEEY---EL 136
DB 145 NLNENFLCESGGCRTHHCEWQKRWYTSERYLESTPLC---YVVASLDDDEGESCEF 201
QY 137 WTTD 140
DB 202 WLSE 205

RESULT 3
Q8IL35 PRELIMINARY; PRT; 1898 AA.
ID 08IL35
AC 08IL35:
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical protein, conserved.
GN Pf14_0414.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-3D7;
MDLINE=22255705; Pubmed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlson J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Sub B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RL Nature 419:498-511(2002).

```

```

DR EMBL; AE014823; AAN37027.1; -.
RW Hypothetical protein.
SQ SEQUENCE 1898 AA; 223829 MW; 7B5CB02A7091822F CRC64;

Query Match 11.2%; Score 104.5; DB 5; Length 1898;
Best Local Similarity 27.5%; Pred. No. 3;
Matches 44; Conservative 23; Mismatches 56; Indels 37; Gaps 9;

QY 17 AMKSLKADVENYVWKATYKN-----DPVGNDFTCGVAMNDVNEDEKSIQAEFL 69
DB 571 AMKSSVIVER---YCKSKYENKKDKLLIDYDMSLFT--HYSN--NHDSKEQSNFF 623
QY 70 MNNA-----DTNQFATEKYAVKMGYN--ENAFRYETEDQVFTDVI-----AY 114
DB 624 HNNANSYTFEDTJSENF-----ISKYFPRISYKFSYDFEEGKFDNLIGCMCY 676
QY 115 SDQNCVDIYVPGTDGNEGEYELWTTVDNIPANLKNFNE 154
DB 677 IYDN-KIVLNDYFKIOTLFKKRYDYDNIIFVCINNPRQ 715

```

```

RESULT 4
Q59983 PRELIMINARY; PRT; 1590 AA.
ID Q59983
AC Q59983:
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-NOV-1996 (Tremblrel. 01, Last annotation update)
DE Glucosyltransferase-I precursor (Ec 2.4.1.5).
GN GTFI.
OS Streptococcus sobrinus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1310;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-OM2176;
MDLINE=94146405; Pubmed=8312602;
RA Sato S., Inoue M., Handa N., Aizawa Y., Isobe Y., Katayama T.;
RT "DNA sequence of the glucosyltransferase gene of serotype d
RT Streptococcus sobrinus."
RL DNA Seq. 4:19-27(1993).
DR EMBL; D13858; BAA02876.1; -.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro-70.
DR Pfam; PF01473; CW_binding_1; 16.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Glucosyltransferase; signal; transferase.
FT SIGNAL 1
FT CHAIN 39 1590
FT SEQUENCE 1590 AA; 175955 MW; C3C83A57CE3C2B0E CRC64;

Query Match 11.0%; Score 103; DB 2; Length 1590;
Best Local Similarity 25.9%; Pred. No. 3.2;
Matches 52; Conservative 19; Mismatches 54; Indels 76; Gaps 12;

QY 2 QPDM-----ADEAANGAHQ--DAWKSILKADVENYVYV-----KATYK-N 38
DB 335 QPQNGSEKRPYDHDHNGALFLFNGQFDLPDQSNRYLRPTNQTGSLDSFTYPPN 394
QY 39 DPVWGNDFTCGVAMNDVNEDEKSIQAEFL-----FNANNADTNMOPATEKYTA 86
DB 395 DPLGGDYF-----LLANVDNSNPVQAEQLMWLHYLLNFGSITANDADAN--FDSIRVDA 448
QY 87 VKMYGYRENAFRFETEDQVFTDVIAYSNDNDVITYPTDNGEEG-----YELMT- 138
DB 449 V-----DAVDADLDIOISSDYLKAAI--GIDNNKNNKANNHVSIVLEAMSD 489
QY 139 -----TDYDNIPANLKNF 152
DB 490 NDTPYLLHDDGDNL-MNMDNKF 509

```

RESULT 5  
ID 055263 PRELIMINARY: PRT: 1590 AA.  
AC 055263;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
DE GTF-1.  
GN GLUCOSYLTRANSFERASE.  
OS Streptococcus sobrinus.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1310;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 33478;  
RA Sato S.;  
RT "DNA and amino-acid sequences of water-insoluble-glucan synthetase  
RT produced from Streptococcus sobrinus ATCC 33478."  
RL Am. Kagoshima Univ. Dental School 16:23-29(1996).  
DR EMBL: D63570; BAA09792.1; -  
DR InterPro: IPR002479; CW\_binding.  
DR InterPro: IPR003318; Glyco\_hydro\_70.  
DR Pfam: PF01473; CW\_binding\_1; 15.  
DR Pfam: PF02324; Glyco\_hydro\_70; 1.  
SQ SEQUENCE 1590 AA: 176058 MW: 9DF7A3F2C6E4FD43 CRC64:

Query Match 10.48; Score 97; DB 2; Length 1590;  
Best Local Similarity 24.98; Pred. No. 11;  
Matches 50; Conservative 22; Mismatches 54; Indels 76; Gaps 12;

QY 2 QPQW-----ADEANGAHQ-DAMKSLKADVENYVYV-----KATYK-N 38  
DB 335 QPQWNGSEKRYDHLQNGALKFPNOSDLPTQSNRLNRTPTNGSLDSFTYNNAN 394  
QY 39 DPVANGDFTCGVYVANDVNEDEKSIQAEFL-----FNNAADTNMQFATEKYTA 86  
DB 395 DPLGGEYF-----LLANDVDSNPVYQAEQLMHLHYLLNFGSIVAKDADAN--FDSIRYDA 448  
QY 87 VKMVGYNRENAFREYEDGVFTVIVASDNCVIVYPRGDGEEG-----YELMT- 138  
DB 449 V-----DNDVADLLQISSDYLAAY--GIDKNNKANNHVSIVEAMSD 489  
QY 139 -----TDYDNIPANCLNKF 152  
DB 490 NDTPYLHDDGDNL-MNMDNKF 509

RESULT 6  
ID 09GS27 PRELIMINARY: PRT: 1449 AA.  
AC 09GS27;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Carbamoyl phosphate synthetase II (EC 1.4.3.6) (Copper amine oxidase)  
DE (Fragment).  
OS Plasmodium berghei.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5821;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K173;  
RA Karib M.T., Davies N.P., Stewart T.S.;  
RT "Unique insert sequences within Carbamoyl Phosphate Synthetase II Gene  
RT from Plasmodium berghei and Plasmodium chabaudi."  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + O(2) = RCHO + NH(3) +  
CC H(2)O(2).  
CC -1- COPACITOR: BINDS 1 COPPER ION AND 1 TOPAQUINONE PER SUBUNIT (BY  
CC SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE COPPER/TOPAQUINONE OXIDASE FAMILY.  
DR EMBL: AF286897; AAC10785.1; -.

DR HSSP; P00968; 1CS0.  
DR InterPro: IPR005483; CPase\_L.  
DR InterPro: IPR005479; CPase\_L\_D2.  
DR InterPro: IPR005480; CPase\_L\_D3.  
DR InterPro: IPR005481; CPase\_L\_N.  
DR InterPro: IPR001317; CPS\_GATase.  
DR InterPro: IPR000269; CUNH\_oxidase.  
DR InterPro: IPR000991; GATase.1.  
DR Pfam: PF00289; CPase\_L\_chain; 1.  
DR Pfam: PF02786; CPase\_L\_D2; 1.  
DR Pfam: PF02787; CPase\_L\_D3; 1.  
DR Pfam: PF00117; GATase; 1.  
DR PRINTS: PR00098; CPASE.  
DR PRINTS: PR00099; CPASGATASE.  
DR PROSITE: PS01164; COPPER\_AMINE\_OXID\_1; 1.  
DR PROSITE: PS00867; CPASE\_2; 1.  
KW Copper; Oxidoreductase; TPO.  
FT NON\_TER 1  
FT TER 1  
SQ SEQUENCE 1449 AA: 167415 MW: CC72C27C72077F6B CRC64:

Query Match 10.38; Score 96.5; DB 5; Length 1449;  
Best Local Similarity 19.68; Pred. No. 11;  
Matches 37; Conservative 36; Mismatches 65; Indels 51; Gaps 9;

QY 6 ADEANGAHQDAMKSLKAD-----VENYVYVYKATYKNDPVNGNDFTCGVYVANDV 56  
DB 1202 SEDEENVNPD--KSMRRDDNRGINNIANNSTYI-----RDSVYNNEXK-INKMRELI 1252  
QY 57 NEDEK--IQAEFLFNNAADTNMQFATEKYVAVMYGYNRENAFREYEDGVFTD--- 110  
DB 1253 NNDKSEVYSKSKFANSTNSKCEYINSNKEAKKINVDNTWYNTNDEEENKNSLS 1312  
QY 111 -----VIASDNN-----CDVIYVPTDGN---EGCYELMTTQYVNIPIANC 148  
DB 1313 NNNEMSKNNNNLINSYNDNKSDFSELSYIQHKKRNINDIENDY-----GYEDLYSGS 1367  
QY 149 LKRFNEYAV 157  
DB 1368 DNCYSTYSI 1376

RESULT 7  
ID 09HE41 PRELIMINARY: PRT: 944 AA.  
AC 09HE41;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
DE Related to importin beta-2 subunit (Transportin).  
GN B9B15.030.  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariales; Sordariaceae; Neurospora.  
OX NCBI\_TaxID=5141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Schulte U., Aign V., Hohnel J., Brandt P., Fartmann B., Holland R.,  
RA Nyakatura G., Mewes H.W., Mannhaupt G.;  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA German Neurospora genome project;  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL451014; CAC18173.2; -.  
DR HSSP; Q92973; IOBK.  
DR InterPro: IPR001494; Importinb\_N.  
DR PROSITE: PS50166; IMPORTIN\_B\_NT; 1.  
SQ SEQUENCE 944 AA: 104908 MW: B5C48DCAE470D5EF CRC64:

Query Match 9.98; Score 92.5; DB 3; Length 944;  
Best Local Similarity 23.28; Pred. No. 14;

Matches 41; Conservative 29; Mismatches 66; Indels 41; Gaps 7;

QY 4 DMADEAN-----GAHODAMSLKADVENYVMKATYKNDPVNGNDFTCVGMANDVNE 58  
 ID 081979 PRELIMINARY; PRT; 176 AA.  
 AC 081979;  
 DT 01-MAR-2003 (Tremblrel. 23, Created)  
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE TSGP4  
 OS Ornithodoros savignyi.  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Argasidae; Ornithodoros.  
 NCBI\_TaxID=69826;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Salivary gland;  
 RX MEDLINE=21317941; PubMed=11425229;  
 RA Mans B.J., Venter J.D., Vrey P.J., Louw A.I., Nelitz A.W.;  
 RT "Identification of putative proteins involved in granule biogenesis of  
 tick salivary glands."  
 RL Electrophoresis 22:1739-1746(2001).  
 DR EMBL: AF452891; AAN76831.1;  
 SQ SEQUENCE 176 AA; 19375 MW; 44B7559F3BCB6A25 CRC64;

Query Match 9.88; Score 92; DB 5; Length 176;  
 Best Local Similarity 25.4%; Pred. No. 1.9;  
 Matches 45; Conservative 23; Mismatches 67; Indels 42; Gaps 12;

QY 4 DMADEANGAHODAMSLKADVENYVMKATYKNDPVNGNDFTCVGMANDVNEDEKST 63  
 ID 095W25 PRELIMINARY; PRT; 221 AA.  
 AC 095W25;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
 DE Histamine binding protein.  
 GN HBP.  
 OS Ixodes scapularis (Black-legged tick) (Deer tick).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Ixodidae; Ixodes.  
 NCBI\_TaxID=6945;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Das S., Banerjee G., Deponte K., Marcantonio N., Kantor F.S.,  
 RA Fikrig E.;  
 RT "Salp25d, an Ixodes scapularis antioxidant, is one of 14

RT Immunodominant antigens in engorged tick salivary glands."  
 RL J. Infect. Dis. 184:0-0(2001).  
 DR EMBL: AF209913; AAK97816.1;  
 SQ SEQUENCE 221 AA; 25359 MW; 7B995ACDC0AFA252 CRC64;

Query Match 9.88; Score 91.5; DB 5; Length 221;  
 Best Local Similarity 24.2%; Pred. No. 2.8;  
 Matches 32; Conservative 24; Mismatches 69; Indels 7; Gaps 5;

QY 45 DFTCVGMANDVNEDEKSLAEFLFMNADTNQFAF-EKVTAVKMG-YNRENAPRYET 102  
 ID 08XP49 PRELIMINARY; PRT; 386 AA.  
 AC 08XP49;  
 DT 01-MAR-2002 (Tremblrel. 20, Created)  
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
 DE Hypothetical protein CPE0116.  
 GN CPE0116.  
 OS Clostridium perfringens.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 NCBI\_TaxID=1502;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=13 / Type A;  
 RX PubMed=11792842;  
 RA Shimizu T., Ohnani K., Hirakawa H., Ohshima K., Yamashita A.,  
 RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;  
 RT "Complete genome sequence of Clostridium perfringens, an anaerobic  
 flesh-eater."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).  
 DR EMBL: AP003185; BAB79822.1;  
 KM Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 386 AA; 44316 MW; F06209C2465839D6 CRC64;

Query Match 9.88; Score 91.5; DB 16; Length 386;  
 Best Local Similarity 22.2%; Pred. No. 5.6;  
 Matches 41; Conservative 32; Mismatches 71; Indels 41; Gaps 7;

QY 14 HODAMSLKADVENYVMKATY-KNDPVWG-----NDFTCVGMANDVNEDEKST 63  
 ID 095W25 PRELIMINARY; PRT; 221 AA.  
 AC 095W25;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
 DE Histamine binding protein.  
 GN HBP.  
 OS Ixodes scapularis (Black-legged tick) (Deer tick).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Ixodidae; Ixodes.  
 NCBI\_TaxID=6945;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Das S., Banerjee G., Deponte K., Marcantonio N., Kantor F.S.,  
 RA Fikrig E.;  
 RT "Salp25d, an Ixodes scapularis antioxidant, is one of 14

QY 112 IAVSDNCDVIYVPGDNGEGY--ELMTDYDNT-----PANCLKNENEAVGRE 160  
 ID 095W25 PRELIMINARY; PRT; 221 AA.  
 AC 095W25;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
 DE Histamine binding protein.  
 GN HBP.  
 OS Ixodes scapularis (Black-legged tick) (Deer tick).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Ixodidae; Ixodes.  
 NCBI\_TaxID=6945;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Das S., Banerjee G., Deponte K., Marcantonio N., Kantor F.S.,  
 RA Fikrig E.;  
 RT "Salp25d, an Ixodes scapularis antioxidant, is one of 14

QY 161 TRDYF 165  
 ID 095W25 PRELIMINARY; PRT; 221 AA.  
 AC 095W25;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
 DE Histamine binding protein.  
 GN HBP.  
 OS Ixodes scapularis (Black-legged tick) (Deer tick).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Ixodidae; Ixodes.  
 NCBI\_TaxID=6945;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Das S., Banerjee G., Deponte K., Marcantonio N., Kantor F.S.,  
 RA Fikrig E.;  
 RT "Salp25d, an Ixodes scapularis antioxidant, is one of 14

QY 329 RNEYI 333  
 ID 095W25 PRELIMINARY; PRT; 221 AA.  
 AC 095W25;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
 DE Histamine binding protein.  
 GN HBP.  
 OS Ixodes scapularis (Black-legged tick) (Deer tick).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Ixodidae; Ixodes.  
 NCBI\_TaxID=6945;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Das S., Banerjee G., Deponte K., Marcantonio N., Kantor F.S.,  
 RA Fikrig E.;  
 RT "Salp25d, an Ixodes scapularis antioxidant, is one of 14

DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
DE NMDH dehydrogenase subunit F (Fragment).  
GN NDMH.  
OS Lythamchia nemorum (yellow pimperl).  
OC Chloroplast.  
OC Eukaryote: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;  
OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:  
OC Asteridae: Ericales; Myrsinaceae; Lysimachia.  
OX NCBI\_TaxID=110756;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kallersjö M., Bergqvist G., Anderberg A.A.;  
RT "Genetic reassignment in primuloid families of the Ericales s. l.  
(Angiosperms): a phylogenetic analysis based on DNA sequences from  
RT three chloroplast genes and morphology.";  
RL Am. J. Bot. 0:0-0(2000).  
DR EMBL: AF213747; AAC43878.1; -  
DR InterPro: IPR003916; NADHub\_oxrd5.  
DR InterPro: IPR002128; Oxidored\_q1.  
DR InterPro: IPR001516; Oxidored\_q1\_C.  
DR Pfam: PF00361; Oxidored\_q1\_1.  
DR Pfam: PF01010; Oxidored\_q1\_C\_1.  
DR Pfam: PF00662; Oxidored\_q1\_N\_1.  
DR PRINTS: PR01434; NADHGNASE5.  
KM NMD; Oxidoreductase; Plastocyanine; Chloroplast.  
FT NON\_TER 1  
FT NON\_TER 1  
SQ SEQUENCE 640 AA; 72317 MW; 120FA9DAE50E09E5 CRC64;

Query Match 9.8%; Score 91.5; DB 8; Length 640;  
Best Local Similarity 21.9%; Pred. No. 10;

Matches 33; Conservative 23; Mismatches 58; Indels 35; Gaps 5;

QY 27 NVY---YVVKATYKNDPFGNDFTCVGMANDVNEDEKSIQAEFLFMNN----- 72  
DB 421 NVYFKNNNGKTSHSISLMK-----GNSKKKKONIFLLNNNESTLKKRI 468  
QY 73 --ADTNKQFATEKVTAVKMYGNRENAFRYETEDGQVFTDIYASDNCVITYPGTGN 130  
DB 469 YRIDENVKQMTFRFTTSTFGNKTYSYPESDWTMLFTLLIFLFTSC--VGFIGIPFN 526  
QY 131 EEGYEL-----WTTDVNIPANCLNKRENYA 156  
DB 527 QEGTDLDILSKMLTPSINLHONLKNYDWS 557

## RESULT 12

QY 08MVC3 PRELIMINARY; PRT: 306 AA.  
AC 08MVC3;  
DT 01-OCT-2002 (TREMBlrel. 22, Created)  
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
DE Putative secreted histamine binding protein.  
OS Ixodes scapularis (Black-legged tick) (Deer tick).  
OC Eukaryote: Metazoa: Arthropoda: Chelicerata: Arachnida; Acari;  
OC Parasitiformes; Ixodida; Ixodidae; Ixodes.  
OX NCBI\_TaxID=6945;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN-Rhode Island; TISSUE-Salivary gland;  
RA Valenzuela J.G., Francischetti I.M., Pham V.M., Garfield M.,  
RA Mather T.N., Ribeiro J.M.C.;  
RT "Exploring the Salivome of the Tick Vector of Lyme Disease, Ixodes  
RT scapularis."  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF483717; AAM93639.1; -  
SQ SEQUENCE 306 AA; 34232 MW; B2DB09181CC4395D CRC64;  
Query Match 9.7%; Score 91; DB 5; Length 306;

Best Local Similarity 24.1%; Pred. No. 4.6;  
Matches 35; Conservative 21; Mismatches 65; Indels 24; Gaps 4;

QY 7 DEANGAHQDAWMSLKADVENYVYKATYKNDPFGNDFTCVGMANDVNEDEKSIQAE 66  
DB 34 NEPSLGLQSAWKAINKSKDDPFVLMFRSRNHEP-----NITCVVVTSLNETLKIYFT 89  
QY 67 FLFPNNADTNMVFATEKVTAVKMYGNRENAFR-----YETEDGQVFTDIYASDD 117  
DB 90 RTYINETGKNIDLTLEVQVRLNQTQYKLENYIRAGLGTSPDKPTPLGS--NMITYGDY 147  
QY 118 NCDIYVPGTD-----GNEEG 133  
DB 148 SCNTSSKPLDMLKAADAVGSEAG 172

## RESULT 13

QY 09PQ49 PRELIMINARY; PRT: 450 AA.  
AC 09PQ49;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Hypothetical protein U0441.  
GN U0441.  
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).  
OC Bacteria: Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.  
OX NCBI\_TaxID=134821;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Serovar 3;  
RX MEDLINE-20500219; PubMed-11048724;  
RA Glass J.I., Lefkowitz E.J., Glass J.S., Helner C.R., Chen E.Y.,  
RA Cassell G.H.;  
RT "The complete sequence of the mucosal pathogen Ureaplasma  
RT urealyticum."  
RL Nature 407:757-762(2000).  
DR EMBL: AEO02140; AAF30853.1; -  
DR InterPro: IPR000209; Peptidase\_S8.  
DR PROSITE: PS00138; SUBTILASE\_S8; 1.  
KM Hypothetical protein; Complete proteome.  
SQ SEQUENCE 450 AA; 52947 MW; 8D0BAB5D1FBA6206 CRC64;  
Query Match 9.7%; Score 91; DB 16; Length 450;  
Best Local Similarity 23.4%; Pred. No. 7.5;  
Matches 37; Conservative 23; Mismatches 60; Indels 38; Gaps 8;  
QY 7 DEANGAHQDAWMSLKADVENYVYKATYKNDPFGNDFTCVGMANDVNEDEKSIQAE 62  
DB 19 DETSENHRIKIKRLDLQNDVKIINHSGWSP---NDPLNIGFYNPINKEIN-DDKD 74  
QY 63 IQAEFL-----FMNNAADTNMVFATEKVTAVKMY-----GYRENAFRYE 101  
DB 75 FKRELKRYIDTLNKKDIDVNTINKKFNKYIDFISYFDIINIFSGNSYND--IRYR 132  
QY 102 TEGDQV---FTDIYAS---DNCVITYPGTGNRE 132  
DB 133 NKNGHILEINWFDISKTSFMDRQNQSIYVGSTKNNE 170  
RESULT 14  
QY 08I3G1 PRELIMINARY; PRT: 1083 AA.  
AC 08I3G1;  
DT 01-MAR-2003 (TREMBlrel. 23, Created)  
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Hypothetical protein.  
GN PF0133W.  
OS Plasmodium falciparum (Isolate 3D7).  
OC Eukaryote: Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=36329;  
RN [1]

RP SEQUENCE FROM N.A.  
RA Devlin K., Baker S., Davies P., Mungai K., Berriman M., Pain A.,  
RA Hall N., Bowman S., Churcher C., Quail M., Barrell B.,  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RN [2]

RP SEQUENCE FROM N.A.  
RA MEDLINE-22253706; PubMed-12368867;  
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,  
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,  
RA Bucke C.O., Burrows C., Cherevach I., Chillingworth C.,  
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Cotton C.,  
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,  
RA Feltwell T., Godle A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,  
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,  
RA Humphray S., Jagels K., James K.D., Johnson D., Kethornou A.,  
RA Knights A., Kontorov B., Kyes S., Larke N., Lawson D., Lennard N.,  
RA Line A., Maddison M., Mclean J., Mooney P., Mule S., Murphy L.,  
RA Oliver K., Ormond D., Price C., Quail M.A., Rabinovitch E.,  
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,  
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,  
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,  
RA Sulston J.E., Craig A., Newbold C., Barrell B.G.;  
RT "Sequence of plasmodium falciparum chromosomes 1, 3-9 and 13.",  
RL Nature 419:527-531(2002).  
DR EMBL: AL929354; CAD51668.1; -  
KM Hypothetical protein.  
SQ SEQUENCE 1083 AA; 128123 MW; 487F49816957C623 CRC64;

Query Match 9.7%; Score 90.5; DB 5; Length 1083;  
Best Local Similarity 23.5%; Pred. No. 25;  
Matches 36; Conservative 25; Mismatches 49; Indels 43; Gaps 7;

QY 25 VENVYVVKATYKNDPVWGNDFTC-----VGYMANDVNEDEKSIQAEFL-FMN-----N 72  
DB 506 INMMFEIKIHNSNDKTKKFCSEYILISINILINKIIECNTTYENYILEFLDLFEKN 565  
QY 73 ADNNMOFATEKVA-VKMTYENRENAFR-----YETEDGOVFTDVIAYSDDNCVT 122  
DB 566 ACVYSKDNTEKRYIEKYKRRKNCFENSYKWDIEFYNIED-----LYINDSKDVI 617  
QY 123 YVPGTDGNEEGYELMTTDDYDINPANCINKNEY 155  
DB 618 YINGKEKE-----GANRTNYIDKY 637

RESULT 15  
Q9P378 PRELIMINARY; PRT; 542 AA.  
ID 09P378;  
AC 09P378;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
DE Protein similar to yeast gas1 glycopospholipid-anchored surface  
DE glycoprotein, putative glycosidase that regulates the crosslinking of  
DE beta-1,6-glucans in the cell wall.  
GN SPAC19B12.02C.  
OS Schizosaccharomyces pombe (fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972h-;  
RA Wood V., Rajandream M.A., Barrell B.G., Seeger K., Harris D.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL390814; CAC000550.1; -  
DR GeneDB\_Spombes; SPAC19B12.02C; -  
DR InterPro: IPR004886; GAS1.  
DR Pfam: PF03198; GAS1; 1.  
SQ SEQUENCE 542 AA; 58115 MW; 1D243FDA5AD2EC3A CRC64;

Query Match 9.6%; Score 89.5; DB 3; Length 542;

Best Local Similarity 25.1%; Pred. No. 13;  
Matches 44; Conservative 22; Mismatches 60; Indels 49; Gaps 10;

QY 13 AHQDAKSLKADYENYVYKATYKNDPVWGNDF-----TCVGMANDVN 57  
DB 102 AFQDAGIYVLSLAQPY---EALSSDPTVTVDLFSRYTEVVDLSLADYDNLGFIAG--N 156  
QY 58 EDEKSIQAEFLFMNADTN-MOFATEKRYAVKMY---GYNRNNAFRYETEDGOVFTDVI 112  
DB 157 E-----VIGNNTNTNAAFYKAVRDYKSYKSSGY-RQIPVGYSTNDEEYTRDPM 206  
QY 113 AY-----SDNCDVI-----YVPGTDGNEEGYELMTTDDYDINPANCINKNEY 155  
DB 207 AYVFDCGDDDDHVDYFGINITYEWCQSDSFVSSGYQERTTEFSMTVPMI--FSEF 259

Search completed: August 1, 2003, 12:19:59  
Job time : 51.1829 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using sw model

Run on: August 1, 2003, 12:15:03 ; Search time 10.4229 Seconds  
(without alignments)  
771.532 Million cell updates/sec

Title: US-10-087-195-7  
Perfect score: 936  
Sequence: 1 NOPWMADEANGAHQDAWKS.....FNEYAVGRETDFTSACLE 171

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues  
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	936	100.0	190	1	HBP2_RHIAP
2	630	67.3	190	1	HBP1_RHIAP
3	331.5	35.4	200	1	HBP2_RHIAP
4	105	11.2	1592	1	GFE2_STRO
5	91	9.7	1597	1	GFE1_STRO
6	89.5	9.6	542	1	YLM2_SCHPO
7	84.5	9.0	858	1	CYAG_DICDI
8	84	9.0	624	1	AMYG_ARXAD
9	83.5	8.9	404	1	CR99_HUMAN
10	83.5	8.8	2911	1	FN22_HUMAN
11	82.5	8.8	954	1	XYNA_RUMFL
12	81.5	8.7	1129	1	RPA2_DROME
13	80	8.5	1695	1	CICA_PAEPP
14	79.5	8.5	2907	1	FN22_MOUSE
15	78.5	8.4	1822	1	ITB4_HUMAN
16	77.5	8.3	421	1	DHE2_PEPAS
17	77.5	8.3	512	1	HEX_ADE09
18	77.5	8.3	1026	1	BGAL_STRTA
19	77	8.2	335	1	Y687_METTA
20	77	8.2	365	1	VG13_BP803
21	76	8.1	393	1	VPAP_HSV6U
22	76	8.1	489	1	INGR_HUMAN
23	75.5	8.1	1279	1	ABU_THESA
24	75.5	8.1	265	1	YR44_MYCCA
25	75	8.0	557	1	PGWU_EMEHI
26	75	8.0	517	1	CD66_HALN1
27	74	7.9	517	1	GUWA_CLOLO
28	73.5	7.9	541	1	YA97_MYCPN
29	73.5	7.9	828	1	PMFC_PROMI
30	73	7.8	625	1	TBP2_HAEIN
31	73	7.8	749	1	SPOT_SPICI
32	73	7.8	960	1	DLG1_DROME
33	73	7.8	960	1	DLG1_DROME

34	72.5	7.7	282	1	POR2_XENIA
35	72.5	7.7	661	1	PSO2_YEAST
36	72.5	7.7	846	1	Y661_MYCTU
37	72.5	7.7	889	1	IRE1_RAT
38	72	7.7	283	1	PHUD_BACE
39	72	7.7	363	1	OMPF_SALTI
40	72	7.7	600	1	ABRA_PLAEG
41	72	7.7	603	1	BGLR_ECOLI
42	72	7.7	743	1	ABRA_PLAEC
43	72	7.7	1041	1	UN83_CAEL
44	72	7.7	1227	1	CJBE_BACTU
45	71.5	7.6	506	1	CYSP_PLAVN

## ALIGNMENTS

RESULT 1					
ID	HBP2_RHIAP	STANDARD:	PRT:	190 AA.	
AC	077421:				
DT	30-MAY-2000 (Rel. 39, Last sequence update)				
DT	30-MAY-2000 (Rel. 39, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Female-specific histamine-binding protein 2 precursor (FS-HBP2).				
OS	Rhipicephalus appendiculatus (Brown ear tick).				
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;				
OC	Parasitiformes; Ixodida; Ixodidae; Rhipicephalus.				
OX	NCBI_TaxID=34631;				
RM	[1]				
RP	SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (1.25 ANGSTROMS).				
RC	TISSUE-Salivary gland;				
RA	MEDLINE=99286454; PubMed=10360182;				
RT	Paesen G.C., Adams P.L., Harlos K., Nuttall P.A., Stuart D.I.;				
RT	"Tick histamine-binding proteins: Isolation, cloning, and three-				
RL	dimensional structure.";				
CC	Mol. Cell 3:661-671(1999).				
CC	-1- FUNCTION: BINDS HISTAMINE WITH A HIGH-AFFINITY. THE ABILITY TO				
CC	OUTCOMPETE HISTAMINE RECEPTORS INDICATES THAT ITS FUNCTION IS TO				
CC	SUPPRESS INFLAMMATION DURING BLOOD FEEDING.				
CC	-1- SUBCELLULAR LOCATION: Secreted.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).				
CC	-----				
DR	EMBL: U96081; AAC63107.1; -				
DR	PDB: 1OFT; 19-APR-00.				
DR	PDB: 1OFV; 19-APR-00.				
DR	InterPro: IPR002970; His_binding.				
DR	Pfam: PF02098; His_binding; 1.				
DR	ProDom: PD152455; His_binding; 1.				
FT	Signal: 3D-structure.				
FT	CHAIN	1	19		
FT		20	190		
FT	DISULFID	67	188		
FT	DISULFID	138	167		
FT	TURN	23	24		
FT	HELI	27	30		
FT	HELI	31	33		
FT	HELI	36	41		
FT	TURN	42	45		
FT	STRAND	48	53		
FT	STRAND	54	58		
FT	TURN	59	61		
FT	TURN	62	63		
FT	TURN	64	64		
FT	STRAND	66	76		
FT	STRAND				

## FEMALE-SPECIFIC HISTAMINE-BINDING PROTEIN

```

FT TURN 77 80
FT STRAND 81 89
FT TURN 91 92
FT STRAND 97 106
FT TURN 109 110
FT STRAND 116 121
FT TURN 122 123
FT STRAND 126 135
FT TURN 136 137
FT STRAND 138 143
FT TURN 152 157
FT TURN 160 161
FT HELIX 165 174
FT TURN 175 177
FT STRAND 181 182
FT TURN 186 188
SQ SEQUENCE 190 AA; 21464 MW; 6923A3E902552B6F CRC64;

Query Match
Best Local Similarity 100.0%; Score 936; DB 1; Length 190;
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NOPWMADEAANGAHQDAWKSIAKADVENYVYVATYKNDPVWGNDFTCVGMADVNEDE 60
DB 20 NOPWMADEAANGAHQDAWKSIAKADVENYVYVATYKNDPVWGNDFTCVGMADVNEDE 79
QY 61 KSIOAEFLFNNMADTNMQFATEKYAVKATGYNRENAFRYETEDGQVFTVIAYSDDNCD 120
DB 80 KSIOAEFLFNNMADTNMQFATEKYAVKATGYNRENAFRYETEDGQVFTVIAYSDDNCD 139
QY 121 VIYVPGTGDGEGEYELMTTIDYDNPANCLKFNENYAVGRETRDVFYSACL 171
DB 140 VIYVPGTGDGEGEYELMTTIDYDNPANCLKFNENYAVGRETRDVFYSACL 190

RESULT 2
HBPL_RHIAP
ID HBPL_RHIAP STANDARD; PRT; 190 AA.
AC 077420;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Female-specific histamine-binding protein 1 precursor (FS-HBPL).
OS Rhipicephalus appendiculatus (Brown ear tick).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Rhipicephalus.
ON NCBI_TaxID=34631;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Salivary gland;
RX MEDLINE=99288454; PubMed=10360182;
RA Priesen G.C., Adams P.L., Harlos K., Nuttall P.A., Stuart D.I.;
RT "Tick histamine-binding proteins: isolation, cloning, and three-
dimensional structure."
RL Mol. Cell 3:661-671(1999).
CC -1- FUNCTION: BINDS HISTAMINE WITH A HIGH-AFFINITY. THE ABILITY TO
OUTCOMPLETE HISTAMINE RECEPTORS INDICATES THAT ITS FUNCTION IS TO
SUPPRESS INFLAMMATION DURING BLOOD FEEDING.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U96080; AAC63106.1; -.
CC HSSP; 077421; 10FT.
CC InterPro; IPR002970; His_binding.
CC Pfam; PF02098; His_binding; 1.
CC ProDom; PD152455; His_binding; 1.
CC ProDom; PD152455; His_binding; 1.

```

```

KW Signal.
FT SIGNAL 1 18
FT CHAIN 19 190
FT DISULFID 66 187
FT DISULFID 137 166
SQ SEQUENCE 190 AA; 21370 MW; 855BE151A90053B1 CRC64;

Query Match
Best Local Similarity 66.5%; Score 630; DB 1; Length 190;
Matches 113; Conservative 20; Mismatches 37; Indels 0; Gaps 0;

QY 1 NOPWMADEAANGAHQDAWKSIAKADVENYVYVATYKNDPVWGNDFTCVGMADVNEDE 60
DB 19 DKPWMADEAANGAHQDAWKSIAKADVENYVYVATYKNDPVWGNDFTCVGMADVNEDE 78
QY 61 KSIOAEFLFNNMADTNMQFATEKYAVKATGYNRENAFRYETEDGQVFTVIAYSDDNCD 120
DB 79 KNEWAMFNFNNMADTNMQFATEKYAVKATGYNRENAFRYETEDGQVFTVIAYSDDNCD 138
QY 121 VIYVPGTGDGEGEYELMTTIDYDNPANCLKFNENYAVGRETRDVFYSACL 170
DB 139 VIYVPGTGDGEGEYELMTTIDYDNPANCLKFNENYAVGRETRDVFYSACL 188

RESULT 3
HBPL_RHIAP
ID HBPL_RHIAP STANDARD; PRT; 200 AA.
AC 077422;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Male-specific histamine-binding salivary protein precursor (MS-HBP).
OS Rhipicephalus appendiculatus (Brown ear tick).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Rhipicephalus.
ON NCBI_TaxID=34631;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Salivary gland;
RX MEDLINE=99288454; PubMed=10360182;
RA Priesen G.C., Adams P.L., Harlos K., Nuttall P.A., Stuart D.I.;
RT "Tick histamine-binding proteins: isolation, cloning, and three-
dimensional structure."
RL Mol. Cell 3:661-671(1999).
CC -1- FUNCTION: BINDS HISTAMINE WITH A HIGH-AFFINITY. THE ABILITY TO
OUTCOMPLETE HISTAMINE RECEPTORS INDICATES THAT ITS FUNCTION IS TO
SUPPRESS INFLAMMATION DURING BLOOD FEEDING.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U96082; AAC63108.1; -.
CC HSSP; 077421; 10FT.
CC InterPro; IPR002970; His_binding.
CC Pfam; PF02098; His_binding; 1.
CC ProDom; PD152455; His_binding; 1.
CC Signal.
FT SIGNAL 1 18
FT CHAIN 19 200
FT DISULFID 65 193
FT DISULFID 137 169
SQ SEQUENCE 200 AA; 22851 MW; C46A1C8C6BCAA008 CRC64;

Query Match
Best Local Similarity 35.4%; Score 331.5; DB 1; Length 200;
Matches 42; Conservative 6; Mismatches 22; Indels 0; Gaps 0;

```



CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: M17391; AAC63063.1; "

DR InterPro: IPR002479; CM\_binding.

DR InterPro: IPR003318; Glyco\_hydro\_70.

DR Pfam: PF01473; CM\_binding\_1; 16.

DR Pfam: PF02324; Glyco\_hydro\_70; 1.

KW Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries.

FT SIGNAL 1 38

FT CHAIN 1 38

FT DOMAIN 39 1050 GLUCAN-BINDING (APPROXIMATE).

FT DOMAIN 1099 1597 GLUCAN-BINDING (APPROXIMATE).

FT DOMAIN 1099 1597 1.25 A, 2 B, AND 5 AC REPEATS.

FT REPEAT 1099 1132 AC REPEAT.

FT REPEAT 1163 1213 AC REPEAT.

FT REPEAT 1227 1277 AC REPEAT.

FT REPEAT 1292 1342 AC REPEAT.

FT REPEAT 1352 1399 B REPEAT.

FT REPEAT 1406 1455 B REPEAT.

FT REPEAT 1465 1512 B REPEAT.

FT REPEAT 1519 1568 AC REPEAT.

FT REPEAT 1582 1597 A REPEAT (INCOMPLETE).

FT SEQUENCE 1597 AA; 177080 MW; B9E86A200868798E CRC64;

Query Match 9.78; Score 91; DB 1; Length 1597;

Best Local Similarity 24.4%; Pred. No. 4.9;

Matches 49; Conservative 21; Mismatches 55; Indels 76; Gaps 12;

QY 2 QPDM-----ADEANGAHO-DAMKSLKADYENYYVY-----KATKK-N 38

DB 341 QPDMNGSEKPYDDHQLNGKLFKDNOSDLPDVSRYRLNRPPTNGSLDSFTYNNAN 400

QY 39 DPVAGNDFTCGVYVANDVNEDEKSIOAEFL-----FMNADJTMKOPATEKYTA 86

DB 401 DPLGVEL-----LLANDVDSNPVQAEQLNMLHYLLNFGTYAKADADAN--FDISIRVDA 454

QY 87 YKMYGYRENAFRFETEDGVFTDYLAISDNCDDVLYVPTDGNEEG-----YELWT- 138

DB 455 V-----DNVADDLQISSDYLLKAAV--GIDNNKNKNNHVSIVEAMSD 495

QY 139 -----TDYDNIPANCLNKF 152

DB 496 NDFPYLHDDGDNL-MNMDNKF 515

RESULT 6

YL2\_SCHPO STANDARD; PRT; 542 AA.

AC 09P378; P78853.

DT 15-SEP-2003 (Rel. 42, Created)

DT 15-SEP-2003 (Rel. 42, Last sequence update)

DE 15-SEP-2003 (Rel. 42, Last annotation update)

DE Hypothetical protein C19B12.02c in chromosome I precursor.

GN SPAC19B12.02C.

OS Schizosaccharomyces pombe (fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI\_TaxID=4896;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=972;

RX MEDLINE=21848401; PubMed=11859360;

RA Mood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

RA Soutouris J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,

RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,

RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,

RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,

RA James K., Jones L., Mungall K., Leather S., McDonald S., McLean J.,

RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,

RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,

RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,

RA Stelton J., Simmonds M., Squares R., Squares S., Stevens K.,

RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,

RA Woodward J., Volckner G., Aert R., Robben J., Gysenprez B.,

RA Wellens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,

RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,

RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,

RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,

RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,

RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RA Lucas R., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,

RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,

RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,

RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,

RA Sipakovski G.V., Ussery D., Barrell B.G., Nurse P.;

RT "The genome sequence of Schizosaccharomyces pombe."

RL Nature 415:671-880(2002).

RN [2]

RP SEQUENCE OF 268-542 FROM N.A.

RC STRAIN=PR745;

RX MEDLINE=98162722; PubMed=9501991;

RA Yoshiooka S., Kato K., Nakai K., Okayama H., Nojima H.;

RT "Identification of open reading frames in Schizosaccharomyces pombe

RT cDNAs."

RL DNA Res. 4:363-369(1997).

CC -1- SIMILARITY: Belongs to the GAS1 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: AL390814; CAC00550.1; "

DR GeneDB: Spombe; SPAC19B12.02c; "

DR InterPro: IPR004886; GAS1.

DR Pfam: PF03198; GAS1; 1.

KW Hypothetical protein; Glycoprotein; Signal.

FT SIGNAL 1 19

FT CHAIN 1 19

FT DOMAIN 20 542

FT DOMAIN 434 520

FT CARBOHYD 35 35

FT CARBOHYD 91 91

FT CARBOHYD 161 161

FT CARBOHYD 249 249

FT CARBOHYD 279 279

FT CARBOHYD 406 406

FT CARBOHYD 484 484

FT CARBOHYD 502 502

FT CARBOHYD 509 509

FT CARBOHYD 509 509

FT CONFLICT 268 278

FT CONFLICT 285 293

FT SEQUENCE 542 AA; 58115 MW; 1D243FDA5AD2E3A CRC64;

Query Match 9.6%; Score 89.5; DB 1; Length 542;

Best Local Similarity 25.1%; Pred. No. 1.8;

Matches 44; Conservative 22; Mismatches 60; Indels 49; Gaps 10;

QY 13 AHODAMSLKADYENYYVYKATYKNDPVWGNDF-----TCGVYANDVN 57

DB 102 AFDQACIYVLSLDAQY-----EAISSDPTWTVDLFSRYTEVVDLSLAPYDMLGFIAG--N 156

QY 58 EDEKSLQAELEFNNDNMDTN-MQFATKRYAVKMY-----GVNRRNARYETEGCOVTVDYI 112

DB 157 E-----VIONNTNTNAAAFVKAARVDSYIYSSG--ROIYGYSTNDEEYTRDPM 206

QY 113 AY-----SDNCDVY-----YVPTDGNEEGYELMTDYDIPANCLNKFNEY 155

DB 207 AYTFDGDGDDDDHDFGINTYEMCGSDSVSSGYQRTFEFSMTVPMI--PSEF 259



DR PRINTS: PR00736; GLHYDRLASE15.  
 DR PROSITE; PS00820; GLUCOAMYLASE; 1.  
 KW Hydrolyase; glycosidase; Polysaccharide degradation; glycoprotein;  
 RM Signal.  
 FT CHAIN 1 18 POTENTIAL.  
 FT ACT\_SITE 19 624 GLUCOAMYLASE.  
 FT ACT\_SITE 340 340 CATALYTIC BASE (BY SIMILARITY).  
 FT ACT\_SITE 343 343 GENERAL ACID CATALYST (BY SIMILARITY).  
 FT ACT\_SITE 344 344 INTERACT WITH SUBSTRATES (BY SIMILARITY).  
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 582 582 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SEQUENCE 624 AA; 68980 MW; 74C2CDFB43FE71B4 CRC64;

Query Match 9.0%; Score 84; DB 1; Length 624;  
 Best Local Similarity 21.5%; Pred. No. 6.5;  
 Matches 45; Conservative 22; Mismatches 64; Indels 78; Gaps 9;

QY 10 ANGAHODA-----WKSLADY-----ENYIMWKATYKN 38  
 DB 299 ANGSDQTVREQLYDSKTYQVAFIKDLQFVANSWSPSPDLMEESESAHFYRLVQK 358  
 QY 39 DPVWGNDFCTGVWANDVNEDEKSIQAEFLMNNADTNMOF-----79  
 DB 359 ALLIGADF-----ANDMGHSLSDKLTQASKRSLDLPFPWMSARQLLYXGVPLRGK 412  
 QY 80 -ATEKTYAV--KMGYRENAFREYEDGOVFTDVIASDNCDDYIYVPGTGNESGYEL 136  
 DB 413 YSKYDLSVYGVNMGVANDVFSY-TND-QILATAYGVSTFSLDYVKVANTTSDESGKPL 470  
 QY 137 WTTDYDNPANCLNKFNAYVGRETRDY 165  
 DB 471 -----GIP-----VGRYPEDVY 482

RESULT 9  
 CT99\_HUMAN STANDARD; PRT; 404 AA.  
 AC Q9N0D5;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE Hypothetical protein C20orf99.  
 GN C20ORF99.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Gardner C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Graffam D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Levesathio M.H., Leverisha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConnachie L.J., McKay K., McMurtry A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,

RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swann R.M., Symcote N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vautin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whitlaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilmink L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 "The DNA sequence and comparative analysis of human chromosome 20.";  
 Nature 414:865-871(2001).

CC -1- SIMILARITY: Contains 3 CCHC-type zinc fingers.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AL034548; CAB81631.1;  
 DR Genew; HGNC:16230; C20orf99.  
 DR InterPro; IPR001878; Znf.CCHC.  
 DR Pfam; PF000098; zf-CCHC; 3.  
 DR PRINTS; PR00939; C2HCZNFINGER.  
 DR SMART; SM00343; Znf\_C2HC; 2.  
 DR PROSITE; PS50158; ZF\_C2HC; 2.  
 KW Hypothetical protein; zinc-finger; Repeat.  
 FT ZN\_FING 334 351 CCHC-TYPE 1.  
 FT ZN\_FING 352 369 CCHC-TYPE 2.  
 FT ZN\_FING 371 388 CCHC-TYPE 3.  
 FT DOMAIN 67 71 POLY-GLY.  
 FT DOMAIN 126 134 POLY-ALA.  
 FT DOMAIN 155 159 POLY-ALA.  
 SO SEQUENCE 404 AA; 43618 MW; C3B728090DD27A4B CRC64;

Query Match 8.9%; Score 83.5; DB 1; Length 404;  
 Best Local Similarity 21.8%; Pred. No. 4.3;  
 Matches 37; Conservative 30; Mismatches 64; Indels 39; Gaps 8;

QY 2 QPMDADEAANGAHODAMKSLKADVENYVYVYKATYKN-----PYWGNDFCTGVWANDVN 57  
 DB 138 RPEGAEDPAERPLQDDEPAALAAAGPGKGRFLVRCFQDDEGACPT--RDFVYVALLRSIG 195  
 QY 58 EDEKSIQAEFLFMNADTNMOF-ATEKVTA-VKMYGYRNN-----96  
 DB 196 MDPSDIYAVIQIGSRFEDVFSRSAEKLAFLVYEKKRQEDQEWENFVLGSKSKSLKT 255  
 QY 97 ---AFRETEBDGOVFTDVIASDNCDDYIYV--GTDGNEEGYELMTTDTY 141  
 DB 256 LFTLFRENEVDVE---DIVTWLKRHCDDVLAIPVKYVD---RGIMTGEY 298

RESULT 10  
 FBN2\_HUMAN STANDARD; PRT; 2911 AA.  
 ID FBN2\_HUMAN  
 AC P35556;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Fibrillin 2 precursor.  
 GN FBN2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94165150; PubMed=8120105;  
 RA Zhang H., Apfelroth S.D., Hu W., Davis E.C., Sangulneti C.,  
 RA Bonadio J., Mecham R.P., Ramirez F.;  
 "Structure and expression of fibrillin-2, a novel microfibrillar

RT component preferentially located in elastic matrices.";  
RL J. Cell Biol. 124:855-863(1994).  
RN [2]  
RP SEQUENCE OF 752-1505 FROM N.A.  
RX MEDLINE-91304567; Pubmed-1852206;  
RA Lee B., Godfrey M., Vitale E., Horl H., Mattei M.-G., Sarfarazi M.,  
RT Tadjipouras P., Ramirez F., Hollister D.;  
"Linkage of Marfan syndrome and a phenotypically related disorder to  
RT two different fibrillin genes.";  
RL Nature 352:330-334(1991).  
RN [3]  
RP VARIANTS CCA TYR-1252 AND SER-1433, AND VARIANT ILE-964.  
RX MEDLINE-96083599; Pubmed-7493032;  
RA Putnam E.A., Zhang H., Ramirez F., Mlewicz D.M.;  
RT "Fibrillin-2 (FBN2) mutations result in the Marfan-like disorder,  
RT congenital contractual atachmodactyly.";  
RL Nat. Genet. 11:456-458(1995).  
RN [4]  
RP VARIANTS CCA HIS-1114.  
RX MEDLINE-98407789; Pubmed-9737771;  
RA Babcock D., Gassner C., Francke U., Maslen C.;  
RT "A single mutation that results in an asp-to-his substitution and  
RT partial exon skipping in a family with congenital contractual  
RT atachmodactyly.";  
RL Hum. Genet. 103:22-28(1998).  
RN [5]  
RP VARIANTS CCA PHE-1141 AND TRP-1252.  
RX MEDLINE-20259236; Pubmed-10797416;  
RA Belleh S., Zhou G., Wang M., Der Kaloustian V.M., Pagon R.A.,  
RT Godfrey M.;  
RT "Two novel fibrillin-2 mutations in congenital contractual  
RT atachmodactyly.";  
RL Am. J. Med. Genet. 92:7-12(2000).  
CC -1- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS  
CC THAT BINDS CALCIUM. FIBRILLIN-2-CONTAINING MICROFIBRILS REGULATE  
CC THE EARLY PROCESS OF ELASTIC FIBER ASSEMBLY.  
CC -1- DISEASE: DEFECTS IN FBN2 ARE THE CAUSE OF CONGENITAL CONTRACTURAL  
CC ARACHNOIDACTYLY (CCA) (ALSO KNOWN AS BEALS SYNDROME). CCA IS  
CC PHENOTYPICALLY SIMILAR TO MARFAN SYNDROME, BUT DOES NOT EFFECT THE  
CC AORTA AND THE EYES.  
CC -1- SIMILARITY: Contains 47 EGF-like domains.  
CC -1- SIMILARITY: Contains 7 TGF-beta binding protein (TGFBRP) domains.  
CC -1- DATABASE: NAME=Elastic Fiber Homepage; NOTE=Fibrillin 2 page;  
CC WWW="http://ef.wustl.edu/genes/FBN2.htm".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U03272; AAA18950.1; -;  
DR EMBL: X62009; -; NOT\_ANNOTATED\_CDS.  
DR PIR: A54105; A54105.  
DR HSSP: P35555; IEMN.  
DR Genew: HGNC:3604; FBN2.  
DR MIM: 121050; -;  
DR GO: GO:0005578; C:extracellular matrix; TAS.  
DR GO: GO:0005201; F:extracellular matrix structural constituent; TAS.  
DR GO: GO:0007345; F:embryogenesis and morphogenesis; TAS.  
DR GO: GO:0007397; P:histogenesis and organogenesis; TAS.  
DR InterPro: IPR000152; Asx\_hydroxyl.  
DR InterPro: IPR001881; EGF\_Ca.  
DR InterPro: IPR001438; EGF\_II.  
DR InterPro: IPR006209; EGF\_Like.  
DR InterPro: IPR002212; Fibrill-assoc.  
DR Pfam: PF000683; EGF; 45.  
DR Pfam: PF00683; TB; 9.  
DR PRINTS: PR00010; EGFBL00D.  
DR SMART: SM00179; EGF\_CA; 43.  
DR PROSITE: PS00010; ASX\_HYDROXYL; 43.

DR PROSITE: PS00022; EGF\_1; 2.  
DR PROSITE: PS01186; EGF\_2; 37.  
DR PROSITE: PS01187; EGF\_CA; 42.  
RW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;  
KW Repeat; Signal; Multigene family; Disease mutation; Polymorphism.  
FT CHAIN 1 29 2911  
FT DOMAIN 111 142  
FT DOMAIN 145 176  
FT DOMAIN 176 207  
FT DOMAIN 275 316  
FT DOMAIN 317 358  
FT REPEAT 359 425  
FT DOMAIN 493 533  
FT DOMAIN 534 573  
FT DOMAIN 574 615  
FT DOMAIN 616 656  
FT DOMAIN 657 697  
FT REPEAT 698 766  
FT DOMAIN 767 808  
FT DOMAIN 809 850  
FT DOMAIN 851 890  
FT DOMAIN 954 995  
FT REPEAT 1071 1071  
FT DOMAIN 1072 1113  
FT DOMAIN 1114 1156  
FT DOMAIN 1157 1198  
FT DOMAIN 1199 1240  
FT DOMAIN 1241 1281  
FT DOMAIN 1282 1323  
FT DOMAIN 1324 1365  
FT DOMAIN 1366 1406  
FT DOMAIN 1407 1447  
FT DOMAIN 1448 1489  
FT DOMAIN 1490 1530  
FT DOMAIN 1531 1571  
FT REPEAT 1572 1648  
FT DOMAIN 1649 1690  
FT DOMAIN 1691 1732  
FT REPEAT 1733 1806  
FT DOMAIN 1807 1848  
FT DOMAIN 1849 1890  
FT DOMAIN 1891 1932  
FT DOMAIN 1933 1971  
FT DOMAIN 1972 2014  
FT DOMAIN 2015 2054  
FT DOMAIN 2055 2096  
FT REPEAT 2097 2169  
FT DOMAIN 2170 2211  
FT DOMAIN 2212 2251  
FT DOMAIN 2252 2292  
FT DOMAIN 2293 2336  
FT DOMAIN 2337 2378  
FT REPEAT 2379 2447  
FT DOMAIN 2448 2489  
FT DOMAIN 2490 2530  
FT DOMAIN 2531 2569  
FT DOMAIN 2570 2612  
FT DOMAIN 2613 2652  
FT DOMAIN 2653 2693  
FT DOMAIN 2694 2733  
FT DISULFID 115 124  
FT DISULFID 119 130  
FT DISULFID 132 141  
FT DISULFID 149 159  
FT DISULFID 153 164  
FT DISULFID 166 175  
FT DISULFID 180 190  
FT DISULFID 184 195  
FT DISULFID 197 206  
FT DISULFID 279 291  
FT DISULFID 286 300  
FT DISULFID 302 315  
EGF-LIKE 1.  
EGF-LIKE 2.  
EGF-LIKE 3.  
EGF-LIKE 4.  
EGF-LIKE 5.  
EGF-LIKE 6.  
EGF-LIKE 7.  
EGF-LIKE 8.  
EGF-LIKE 9.  
EGF-LIKE 10.  
EGF-LIKE 11.  
EGF-LIKE 12.  
EGF-LIKE 13.  
EGF-LIKE 14.  
EGF-LIKE 15.  
EGF-LIKE 16.  
EGF-LIKE 17.  
EGF-LIKE 18.  
EGF-LIKE 19.  
EGF-LIKE 20.  
EGF-LIKE 21.  
EGF-LIKE 22.  
EGF-LIKE 23.  
EGF-LIKE 24.  
EGF-LIKE 25.  
EGF-LIKE 26.  
EGF-LIKE 27.  
EGF-LIKE 28.  
EGF-LIKE 29.  
EGF-LIKE 30.  
EGF-LIKE 31.  
EGF-LIKE 32.  
EGF-LIKE 33.  
EGF-LIKE 34.  
EGF-LIKE 35.  
EGF-LIKE 36.  
EGF-LIKE 37.  
EGF-LIKE 38.  
EGF-LIKE 39.  
EGF-LIKE 40.  
EGF-LIKE 41.  
EGF-LIKE 42.  
EGF-LIKE 43.  
EGF-LIKE 44.  
EGF-LIKE 45.  
EGF-LIKE 46.  
EGF-LIKE 47.  
EGF-LIKE 48.  
EGF-LIKE 49.  
EGF-LIKE 50.  
EGF-LIKE 51.  
EGF-LIKE 52.  
EGF-LIKE 53.  
EGF-LIKE 54.  
EGF-LIKE 55.  
EGF-LIKE 56.  
EGF-LIKE 57.  
EGF-LIKE 58.  
EGF-LIKE 59.  
EGF-LIKE 60.  
EGF-LIKE 61.  
EGF-LIKE 62.  
EGF-LIKE 63.  
EGF-LIKE 64.  
EGF-LIKE 65.  
EGF-LIKE 66.  
EGF-LIKE 67.  
EGF-LIKE 68.  
EGF-LIKE 69.  
EGF-LIKE 70.  
EGF-LIKE 71.  
EGF-LIKE 72.  
EGF-LIKE 73.  
EGF-LIKE 74.  
EGF-LIKE 75.  
EGF-LIKE 76.  
EGF-LIKE 77.  
EGF-LIKE 78.  
EGF-LIKE 79.  
EGF-LIKE 80.  
EGF-LIKE 81.  
EGF-LIKE 82.  
EGF-LIKE 83.  
EGF-LIKE 84.  
EGF-LIKE 85.  
EGF-LIKE 86.  
EGF-LIKE 87.  
EGF-LIKE 88.  
EGF-LIKE 89.  
EGF-LIKE 90.  
EGF-LIKE 91.  
EGF-LIKE 92.  
EGF-LIKE 93.  
EGF-LIKE 94.  
EGF-LIKE 95.  
EGF-LIKE 96.  
EGF-LIKE 97.  
EGF-LIKE 98.  
EGF-LIKE 99.  
EGF-LIKE 100.  
EGF-LIKE 101.  
EGF-LIKE 102.  
EGF-LIKE 103.  
EGF-LIKE 104.  
EGF-LIKE 105.  
EGF-LIKE 106.  
EGF-LIKE 107.  
EGF-LIKE 108.  
EGF-LIKE 109.  
EGF-LIKE 110.  
EGF-LIKE 111.  
EGF-LIKE 112.  
EGF-LIKE 113.  
EGF-LIKE 114.  
EGF-LIKE 115.  
EGF-LIKE 116.  
EGF-LIKE 117.  
EGF-LIKE 118.  
EGF-LIKE 119.  
EGF-LIKE 120.  
EGF-LIKE 121.  
EGF-LIKE 122.  
EGF-LIKE 123.  
EGF-LIKE 124.  
EGF-LIKE 125.  
EGF-LIKE 126.  
EGF-LIKE 127.  
EGF-LIKE 128.  
EGF-LIKE 129.  
EGF-LIKE 130.  
EGF-LIKE 131.  
EGF-LIKE 132.  
EGF-LIKE 133.  
EGF-LIKE 134.  
EGF-LIKE 135.  
EGF-LIKE 136.  
EGF-LIKE 137.  
EGF-LIKE 138.  
EGF-LIKE 139.  
EGF-LIKE 140.  
EGF-LIKE 141.  
EGF-LIKE 142.  
EGF-LIKE 143.  
EGF-LIKE 144.  
EGF-LIKE 145.  
EGF-LIKE 146.  
EGF-LIKE 147.  
EGF-LIKE 148.  
EGF-LIKE 149.  
EGF-LIKE 150.  
EGF-LIKE 151.  
EGF-LIKE 152.  
EGF-LIKE 153.  
EGF-LIKE 154.  
EGF-LIKE 155.  
EGF-LIKE 156.  
EGF-LIKE 157.  
EGF-LIKE 158.  
EGF-LIKE 159.  
EGF-LIKE 160.  
EGF-LIKE 161.  
EGF-LIKE 162.  
EGF-LIKE 163.  
EGF-LIKE 164.  
EGF-LIKE 165.  
EGF-LIKE 166.  
EGF-LIKE 167.  
EGF-LIKE 168.  
EGF-LIKE 169.  
EGF-LIKE 170.  
EGF-LIKE 171.  
EGF-LIKE 172.  
EGF-LIKE 173.  
EGF-LIKE 174.  
EGF-LIKE 175.  
EGF-LIKE 176.  
EGF-LIKE 177.  
EGF-LIKE 178.  
EGF-LIKE 179.  
EGF-LIKE 180.  
EGF-LIKE 181.  
EGF-LIKE 182.  
EGF-LIKE 183.  
EGF-LIKE 184.  
EGF-LIKE 185.  
EGF-LIKE 186.  
EGF-LIKE 187.  
EGF-LIKE 188.  
EGF-LIKE 189.  
EGF-LIKE 190.  
EGF-LIKE 191.  
EGF-LIKE 192.  
EGF-LIKE 193.  
EGF-LIKE 194.  
EGF-LIKE 195.  
EGF-LIKE 196.  
EGF-LIKE 197.  
EGF-LIKE 198.  
EGF-LIKE 199.  
EGF-LIKE 200.  
EGF-LIKE 201.  
EGF-LIKE 202.  
EGF-LIKE 203.  
EGF-LIKE 204.  
EGF-LIKE 205.  
EGF-LIKE 206.  
EGF-LIKE 207.  
EGF-LIKE 208.  
EGF-LIKE 209.  
EGF-LIKE 210.  
EGF-LIKE 211.  
EGF-LIKE 212.  
EGF-LIKE 213.  
EGF-LIKE 214.  
EGF-LIKE 215.  
EGF-LIKE 216.  
EGF-LIKE 217.  
EGF-LIKE 218.  
EGF-LIKE 219.  
EGF-LIKE 220.  
EGF-LIKE 221.  
EGF-LIKE 222.  
EGF-LIKE 223.  
EGF-LIKE 224.  
EGF-LIKE 225.  
EGF-LIKE 226.  
EGF-LIKE 227.  
EGF-LIKE 228.  
EGF-LIKE 229.  
EGF-LIKE 230.  
EGF-LIKE 231.  
EGF-LIKE 232.  
EGF-LIKE 233.  
EGF-LIKE 234.  
EGF-LIKE 235.  
EGF-LIKE 236.  
EGF-LIKE 237.  
EGF-LIKE 238.  
EGF-LIKE 239.  
EGF-LIKE 240.  
EGF-LIKE 241.  
EGF-LIKE 242.  
EGF-LIKE 243.  
EGF-LIKE 244.  
EGF-LIKE 245.  
EGF-LIKE 246.  
EGF-LIKE 247.  
EGF-LIKE 248.  
EGF-LIKE 249.  
EGF-LIKE 250.  
EGF-LIKE 251.  
EGF-LIKE 252.  
EGF-LIKE 253.  
EGF-LIKE 254.  
EGF-LIKE 255.  
EGF-LIKE 256.  
EGF-LIKE 257.  
EGF-LIKE 258.  
EGF-LIKE 259.  
EGF-LIKE 260.  
EGF-LIKE 261.  
EGF-LIKE 262.  
EGF-LIKE 263.  
EGF-LIKE 264.  
EGF-LIKE 265.  
EGF-LIKE 266.  
EGF-LIKE 267.  
EGF-LIKE 268.  
EGF-LIKE 269.  
EGF-LIKE 270.  
EGF-LIKE 271.  
EGF-LIKE 272.  
EGF-LIKE 273.  
EGF-LIKE 274.  
EGF-LIKE 275.  
EGF-LIKE 276.  
EGF-LIKE 277.  
EGF-LIKE 278.  
EGF-LIKE 279.  
EGF-LIKE 280.  
EGF-LIKE 281.  
EGF-LIKE 282.  
EGF-LIKE 283.  
EGF-LIKE 284.  
EGF-LIKE 285.  
EGF-LIKE 286.  
EGF-LIKE 287.  
EGF-LIKE 288.  
EGF-LIKE 289.  
EGF-LIKE 290.  
EGF-LIKE 291.  
EGF-LIKE 292.  
EGF-LIKE 293.  
EGF-LIKE 294.  
EGF-LIKE 295.  
EGF-LIKE 296.  
EGF-LIKE 297.  
EGF-LIKE 298.  
EGF-LIKE 299.  
EGF-LIKE 300.  
EGF-LIKE 301.  
EGF-LIKE 302.  
EGF-LIKE 303.  
EGF-LIKE 304.  
EGF-LIKE 305.  
EGF-LIKE 306.  
EGF-LIKE 307.  
EGF-LIKE 308.  
EGF-LIKE 309.  
EGF-LIKE 310.  
EGF-LIKE 311.  
EGF-LIKE 312.  
EGF-LIKE 313.  
EGF-LIKE 314.  
EGF-LIKE 315.  
EGF-LIKE 316.  
EGF-LIKE 317.  
EGF-LIKE 318.  
EGF-LIKE 319.  
EGF-LIKE 320.  
EGF-LIKE 321.  
EGF-LIKE 322.  
EGF-LIKE 323.  
EGF-LIKE 324.  
EGF-LIKE 325.  
EGF-LIKE 326.  
EGF-LIKE 327.  
EGF-LIKE 328.  
EGF-LIKE 329.  
EGF-LIKE 330.  
EGF-LIKE 331.  
EGF-LIKE 332.  
EGF-LIKE 333.  
EGF-LIKE 334.  
EGF-LIKE 335.  
EGF-LIKE 336.  
EGF-LIKE 337.  
EGF-LIKE 338.  
EGF-LIKE 339.  
EGF-LIKE 340.  
EGF-LIKE 341.  
EGF-LIKE 342.  
EGF-LIKE 343.  
EGF-LIKE 344.  
EGF-LIKE 345.  
EGF-LIKE 346.  
EGF-LIKE 347.  
EGF-LIKE 348.  
EGF-LIKE 349.  
EGF-LIKE 350.  
EGF-LIKE 351.  
EGF-LIKE 352.  
EGF-LIKE 353.  
EGF-LIKE 354.  
EGF-LIKE 355.  
EGF-LIKE 356.  
EGF-LIKE 357.  
EGF-LIKE 358.  
EGF-LIKE 359.  
EGF-LIKE 360.  
EGF-LIKE 361.  
EGF-LIKE 362.  
EGF-LIKE 363.  
EGF-LIKE 364.  
EGF-LIKE 365.  
EGF-LIKE 366.  
EGF-LIKE 367.  
EGF-LIKE 368.  
EGF-LIKE 369.  
EGF-LIKE 370.  
EGF-LIKE 371.  
EGF-LIKE 372.  
EGF-LIKE 373.  
EGF-LIKE 374.  
EGF-LIKE 375.  
EGF-LIKE 376.  
EGF-LIKE 377.  
EGF-LIKE 378.  
EGF-LIKE 379.  
EGF-LIKE 380.  
EGF-LIKE 381.  
EGF-LIKE 382.  
EGF-LIKE 383.  
EGF-LIKE 384.  
EGF-LIKE 385.  
EGF-LIKE 386.  
EGF-LIKE 387.  
EGF-LIKE 388.  
EGF-LIKE 389.  
EGF-LIKE 390.  
EGF-LIKE 391.  
EGF-LIKE 392.  
EGF-LIKE 393.  
EGF-LIKE 394.  
EGF-LIKE 395.  
EGF-LIKE 396.  
EGF-LIKE 397.  
EGF-LIKE 398.  
EGF-LIKE 399.  
EGF-LIKE 400.  
EGF-LIKE 401.  
EGF-LIKE 402.  
EGF-LIKE 403.  
EGF-LIKE 404.  
EGF-LIKE 405.  
EGF-LIKE 406.  
EGF-LIKE 407.  
EGF-LIKE 408.  
EGF-LIKE 409.  
EGF-LIKE 410.  
EGF-LIKE 411.  
EGF-LIKE 412.  
EGF-LIKE 413.  
EGF-LIKE 414.  
EGF-LIKE 415.  
EGF-LIKE 416.  
EGF-LIKE 417.  
EGF-LIKE 418.  
EGF-LIKE 419.  
EGF-LIKE 420.  
EGF-LIKE 421.  
EGF-LIKE 422.  
EGF-LIKE 423.  
EGF-LIKE 424.  
EGF-LIKE 425.  
EGF-LIKE 426.  
EGF-LIKE 427.  
EGF-LIKE 428.  
EGF-LIKE 429.  
EGF-LIKE 430.  
EGF-LIKE 431.  
EGF-LIKE 432.  
EGF-LIKE 433.  
EGF-LIKE 434.  
EGF-LIKE 435.  
EGF-LIKE 436.  
EGF-LIKE 437.  
EGF-LIKE 438.  
EGF-LIKE 439.  
EGF-LIKE 440.  
EGF-LIKE 441.  
EGF-LIKE 442.  
EGF-LIKE 443.  
EGF-LIKE 444.  
EGF-LIKE 445.  
EGF-LIKE 446.  
EGF-LIKE 447.  
EGF-LIKE 448.  
EGF-LIKE 449.  
EGF-LIKE 450.  
EGF-LIKE 451.  
EGF-LIKE 452.  
EGF-LIKE 453.  
EGF-LIKE 454.  
EGF-LIKE 455.  
EGF-LIKE 456.  
EGF-LIKE 457.  
EGF-LIKE 458.  
EGF-LIKE 459.  
EGF-LIKE 460.  
EGF-LIKE 461.  
EGF-LIKE 462.  
EGF-LIKE 463.  
EGF-LIKE 464.  
EGF-LIKE 465.  
EGF-LIKE 466.  
EGF-LIKE 467.  
EGF-LIKE 468.  
EGF-LIKE 469.  
EGF-LIKE 470.  
EGF-LIKE 471.  
EGF-LIKE 472.  
EGF-LIKE 473.  
EGF-LIKE 474.  
EGF-LIKE 475.  
EGF-LIKE 476.  
EGF-LIKE 477.  
EGF-LIKE 478.  
EGF-LIKE 479.  
EGF-LIKE 480.  
EGF-LIKE 481.  
EGF-LIKE 482.  
EGF-LIKE 483.  
EGF-LIKE 484.  
EGF-LIKE 485.  
EGF-LIKE 486.  
EGF-LIKE 487.  
EGF-LIKE 488.  
EGF-LIKE 489.  
EGF-LIKE 490.  
EGF-LIKE 491.  
EGF-LIKE 492.  
EGF-LIKE 493.  
EGF-LIKE 494.  
EGF-LIKE 495.  
EGF-LIKE 496.  
EGF-LIKE 497.  
EGF-LIKE 498.  
EGF-LIKE 499.  
EGF-LIKE 500.  
EGF-LIKE 501.  
EGF-LIKE 502.  
EGF-LIKE 503.  
EGF-LIKE 504.  
EGF-LIKE 505.  
EGF-LIKE 506.  
EGF-LIKE 507.  
EGF-LIKE 508.  
EGF-LIKE 509.  
EGF-LIKE 510.  
EGF-LIKE 511.  
EGF-LIKE 512.  
EGF-LIKE 513.  
EGF-LIKE 514.  
EGF-LIKE 515.  
EGF-LIKE 516.  
EGF-LIKE 517.  
EGF-LIKE 518.  
EGF-LIKE 519.  
EGF-LIKE 520.  
EGF-LIKE 521.  
EGF-LIKE 522.  
EGF-LIKE 523.  
EGF-LIKE 524.  
EGF-LIKE 525.  
EGF-LIKE 526.  
EGF-LIKE 527.  
EGF-LIKE 528.  
EGF-LIKE 529.  
EGF-LIKE 530.  
EGF-LIKE 531.  
EGF-LIKE 532.  
EGF-LIKE 533.  
EGF-LIKE 534.  
EGF-LIKE 535.  
EGF-LIKE 536.  
EGF-LIKE 537.  
EGF-LIKE 538.  
EGF-LIKE 539.  
EGF-LIKE 540.  
EGF-LIKE 541.  
EGF-LIKE 542.  
EGF-LIKE 543.  
EGF-LIKE 544.  
EGF-LIKE 545.  
EGF-LIKE 546.  
EGF-LIKE 547.  
EGF-LIKE 548.  
EGF-LIKE 549.  
EGF-LIKE 550.  
EGF-LIKE 551.  
EGF-LIKE 552.  
EGF-LIKE 553.  
EGF-LIKE 554.  
EGF-LIKE 555.  
EGF-LIKE 556.  
EGF-LIKE 557.  
EGF-LIKE 558.  
EGF-LIKE 559.  
EGF-LIKE 560.  
EGF-LIKE 561.  
EGF-LIKE 562.  
EGF-LIKE 563.  
EGF-LIKE 564.  
EGF-LIKE 565.  
EGF-LIKE 566.  
EGF-LIKE 567.  
EGF-LIKE 568.  
EGF-LIKE 569.  
EGF-LIKE 570.  
EGF-LIKE 571.  
EGF-LIKE 572.  
EGF-LIKE 573.  
EGF-LIKE 574.  
EGF-LIKE 575.  
EGF-LIKE 576.  
EGF-LIKE 577.  
EGF-LIKE 578.  
EGF-LIKE 579.  
EGF-LIKE 580.  
EGF-LIKE 581.  
EGF-LIKE 582.  
EGF-LIKE 583.  
EGF-LIKE 584.  
EGF-LIKE 585.  
EGF-LIKE 586.  
EGF-LIKE 587.  
EGF-LIKE 588.  
EGF-LIKE 589.  
EGF-LIKE 590.  
EGF-LIKE 591.  
EGF-LIKE 592.  
EGF-LIKE 593.  
EGF-LIKE 594.  
EGF-LIKE 595.  
EGF-LIKE 596.  
EGF-LIKE 597.  
EGF-LIKE 598.  
EGF-LIKE 599.  
EGF-LIKE 600.  
EGF-LIKE 601.  
EGF-LIKE 602.  
EGF-LIKE 603.  
EGF-LIKE 604.  
EGF-LIKE 605.  
EGF-LIKE 606.  
EGF-LIKE 607.  
EGF-LIKE 608.  
EGF-LIKE 609.  
EGF-LIKE 610.  
EGF-LIKE 611.  
EGF-LIKE 612.  
EGF-LIKE 613.  
EGF-LIKE 614.  
EGF-LIKE 615.  
EGF-LIKE 616.  
EGF-LIKE 617.  
EGF-LIKE 618.  
EGF-LIKE 619.  
EGF-LIKE 620.  
EGF-LIKE 621.  
EGF-LIKE 622.  
EGF-LIKE 623.  
EGF-LIKE 624.  
EGF-LIKE 625.  
EGF-LIKE 626.  
EGF-LIKE 627.  
EGF-LIKE 628.  
EGF-LIKE 629.  
EGF-LIKE 630.  
EGF-LIKE 631.  
EGF-LIKE 632.  
EGF-LIKE 633.  
EGF-LIKE 634.  
EGF-LIKE 635.  
EGF-LIKE 636.  
EGF-LIKE 637.  
EGF-LIKE 638.  
EGF-LIKE 639.  
EGF-LIKE 640.  
EGF-LIKE 641.  
EGF-LIKE 642.  
EGF-LIKE 643.  
EGF-LIKE 644.  
EGF-LIKE 645.  
EGF-LIKE 646.  
EGF-LIKE 647.  
EGF-LIKE 648.  
EGF-LIKE 649.  
EGF-LIKE 650.  
EGF-LIKE 651.  
EGF-LIKE 652.  
EGF-LIKE 653.  
EGF-LIKE 654.  
EGF-LIKE 655.  
EGF-LIKE 656.  
EGF-LIKE 657.  
EGF-LIKE 658.  
EGF-LIKE 659.  
EGF-LIKE 660.  
EGF-LIKE 661.  
EGF-LIKE 662.  
EGF-LIKE 663.  
EGF-LIKE 664.  
EGF-LIKE 665.  
EGF-LIKE 666.  
EGF-LIKE 667.  
EGF-LIKE 668.  
EGF-LIKE 669.  
EGF-LIKE 670.  
EGF-LIKE 671.  
EGF-LIKE 672.  
EGF-LIKE 673.  
EGF-LIKE 674.  
EGF-LIKE 675.  
EGF-LIKE 676.  
EGF-LIKE 677.  
EGF-LIKE 678.  
EGF-LIKE 679.  
EGF-LIKE 680.  
EGF-LIKE 681.  
EGF-LIKE 682.  
EGF-LIKE 683.  
EGF-LIKE 684.  
EGF-LIKE 685.  
EGF-LIKE 686.  
EGF-LIKE 687.  
EGF-LIKE 688.  
EGF-LIKE 689.  
EGF-LIKE 690.  
EGF-LIKE 691.  
EGF-LIKE 692.  
EGF-LIKE 693.  
EGF-LIKE 694.  
EGF-LIKE 695.  
EGF-LIKE 696.  
EGF-LIKE 697.  
EGF-LIKE 698.  
EGF-LIKE 699.  
EGF-LIKE 700.  
EGF-LIKE 701.  
EGF-LIKE 702.  
EGF-LIKE 703.  
EGF-LIKE 704.  
EGF-LIKE 705.  
EGF-LIKE 706.  
EGF-LIKE 707.  
EGF-LIKE 708.  
EGF-LIKE 709.  
EGF-LIKE 710.  
EGF-LIKE 711.  
EGF-LIKE 712.  
EGF-LIKE 713.  
EGF-LIKE 714.  
EGF-LIKE 715.  
EGF-LIKE 716.  
EGF-LIKE 717.  
EGF-LIKE 718.  
EGF-LIKE 719.  
EGF-LIKE 720.  
EGF-LIKE 721.  
EGF-LIKE 722.  
EGF-LIKE 723.  
EGF-LIKE 724.  
EGF-LIKE 725.  
EGF-LIKE 726.  
EGF-LIKE 727.  
EGF-LIKE 728.  
EGF-LIKE 729.  
EGF-LIKE 730.  
EGF-LIKE 731.  
EGF-LIKE 732.  
EGF-LIKE 733.  
EGF-LIKE 734.  
EGF-LIKE 735.  
EGF-LIKE 736.  
EGF-LIKE 737.  
EGF-LIKE 738.  
EGF-LIKE 739.  
EGF-LIKE 740.  
EGF-LIKE 741.  
EGF-LIKE 742.  
EGF-LIKE 743.  
EGF-LIKE 744.  
EGF-LIKE 745.  
EGF-LIKE 746.  
EGF-LIKE 747.  
EGF-LIKE 748.  
EGF-LIKE 749.  
EGF-LIKE 750.  
EGF-LIKE 751.  
EGF-LIKE 752.  
EGF-LIKE 753.  
EGF-LIKE 754.  
EGF-LIKE 755.  
EGF-LIKE 756.  
EGF-LIKE 757.  
EGF-LIKE 758.  
EGF-LIKE 759.  
EGF-LIKE 760.  
EGF-LIKE 761.  
EGF-LIKE 762.  
EGF-LIKE 763.  
EGF-LIKE 764.  
EGF-LIKE 765.  
EGF-LIKE 766.  
EGF-LIKE 767.  
EGF-LIKE 768.  
EGF-LIKE 769.  
EGF-LIKE 770.  
EGF-LIKE 771.  
EGF-LIKE 772.  
EGF-LIKE 773.  
EGF-LIKE 774.  
EGF-LIKE 775.  
EGF-LIKE 776.  
EGF-LIKE 777.  
EGF-LIKE 778.  
EGF-LIKE 779.  
EGF-LIKE 780.  
EGF-LIKE 781.  
EGF-LIKE 782.  
EGF-LIKE 783.  
EGF-LIKE 784.  
EGF-LIKE 785.  
EGF-LIKE 786.  
EGF-LIKE 787.  
EGF-LIKE 788.  
EGF-LIKE 789.  
EGF-LIKE 790.  
EGF-LIKE 791.  
EGF-LIKE 792.  
EGF-LIKE 793.  
EGF-LIKE 794.  
EGF-LIKE 795.  
EGF-LIKE 796.  
EGF-LIKE 797.  
EGF-LIKE 798.  
EGF-LIKE 799.  
EGF-LIKE 800.  
EGF-LIKE 801.  
EGF-LIKE 802.  
EGF-LIKE 803.  
EGF-LIKE 804.  
EGF-LIKE 805.  
EGF-LIKE 806.  
EGF-LIKE 807.  
EGF-LIKE 808.  
EGF-LIKE 809.  
EGF-LIKE 810.  
EGF-LIKE 811.  
EGF-LIKE 812.  
EGF-LIKE 813.  
EGF-LIKE 814.  
EGF-LIKE 815.  
EGF-LIKE 816.  
EGF-LIKE 817.  
EGF-LIKE 818.  
EGF-LIKE 819.  
EGF-LIKE 820.  
EGF-LIKE 821.  
EGF-LIKE 822.  
EGF-LIKE 823.  
EGF-LIKE 824.  
EGF-LIKE 825.  
EGF-LIKE 826.  
EGF-LIKE 827.  
EGF-LIKE 828.  
EGF-LIKE 829.  
EGF-LIKE 830.  
EGF-LIKE 831.  
EGF-LIKE 832.  
EGF-LIKE 833.  
EGF-LIKE 834.  
EGF-LIKE 835.  
EGF-LIKE 836.  
EGF-LIKE 837.  
EGF-LIKE 838.  
EGF-LIKE 839.  
EGF-LIKE 840.  
EGF-LIKE 841.  
EGF-LIKE 842.  
EGF-LIKE 843.  
EGF-LIKE 844.  
EGF-LIKE 845.  
EGF-LIKE 846.  
EGF-LIKE 847.  
EGF-LIKE 848.  
EGF-LIKE 849.  
EGF-LIKE 850.  
EGF-LIKE 851.  
EGF-LIKE 852.  
EGF-LIKE 853.  
EGF-LIKE 854.  
EGF-LIKE 855.  
EGF-LIKE 856.  
EGF-LIKE 857.  
EGF-LIKE 858.  
EGF-LIKE 859.  
EGF-LIKE 860.  
EGF-LIKE 861.  
EGF-LIKE 862.  
EGF-LIKE 863.  
EGF-LIKE 864.  
EGF-LIKE 865.  
EGF-LIKE 866.  
EGF-LIKE 867.  
EGF-LIKE 868.  
EGF-LIKE 869.  
EGF-LIKE 870.  
EGF-LIKE 871.  
EGF-LIKE 872.  
EGF-LIKE 873.  
EGF-LIKE 874.  
EGF-LIKE 875.  
EGF-LIKE 876.  
EGF-LIKE 877.  
EGF-LIKE 878.  
EGF-LIKE 879.  
EGF-LIKE 880.  
EGF-LIKE 881.  
EGF-LIKE 882.  
EGF-LIKE 883.  
EGF-LIKE 884.  
EGF-LIKE 885.  
EGF-LIKE 886.  
EGF-LIKE 887.  
EGF-LIKE 888.  
EGF-LIKE 889.  
EGF-LIKE 890.  
EGF-LIKE 891.  
EGF-LIKE 892.  
EGF-LIKE 893.<

```

FT DISULFID 321 333 BY SIMILARITY.
FT DISULFID 328 342 BY SIMILARITY.
FT DISULFID 344 357 BY SIMILARITY.
FT DISULFID 497 509 BY SIMILARITY.
FT DISULFID 504 518 BY SIMILARITY.
FT DISULFID 520 532 BY SIMILARITY.
FT DISULFID 538 548 BY SIMILARITY.
FT DISULFID 543 557 BY SIMILARITY.
FT DISULFID 559 572 BY SIMILARITY.
FT DISULFID 578 590 BY SIMILARITY.
FT DISULFID 585 599 BY SIMILARITY.
FT DISULFID 601 614 BY SIMILARITY.
FT DISULFID 620 631 BY SIMILARITY.
FT DISULFID 626 640 BY SIMILARITY.
FT DISULFID 642 655 BY SIMILARITY.
FT DISULFID 661 672 BY SIMILARITY.
FT DISULFID 667 681 BY SIMILARITY.
FT DISULFID 683 696 BY SIMILARITY.
FT DISULFID 771 783 BY SIMILARITY.
FT DISULFID 778 792 BY SIMILARITY.
FT DISULFID 794 807 BY SIMILARITY.
FT DISULFID 813 825 BY SIMILARITY.
FT DISULFID 820 834 BY SIMILARITY.
FT DISULFID 836 849 BY SIMILARITY.
FT DISULFID 855 865 BY SIMILARITY.
FT DISULFID 860 874 BY SIMILARITY.
FT DISULFID 876 889 BY SIMILARITY.
FT DISULFID 958 970 BY SIMILARITY.
FT DISULFID 965 979 BY SIMILARITY.
FT DISULFID 981 994 BY SIMILARITY.
FT DISULFID 1076 1088 BY SIMILARITY.
FT DISULFID 1083 1097 BY SIMILARITY.
FT DISULFID 1099 1112 BY SIMILARITY.
FT DISULFID 1118 1130 BY SIMILARITY.
FT DISULFID 1125 1139 BY SIMILARITY.
FT DISULFID 1141 1155 BY SIMILARITY.
FT DISULFID 1161 1173 BY SIMILARITY.

```

Query Match 8.98; Score 83.5; DB 1; Length 2911;  
 Best Local Similarity 29.0%; Pred. No. 44;  
 Matches 49; Conservative 16; Mismatches 45; Indels 59; Gaps 13;

```

QY 7 DEANAGHODAMSLKADVENYVYMKATYK---NDPWGNDPTCYGVANDVNEDEKSI 63
DB 1409 DECSNCTHO---CSINACVN---TPGSYRCACSGFTGDEGTC---SDVDECAENI 1456
QY 64 QAELFMNADPTMOPATEKVTAVKMYGYNRENAFRTEDGOVFT---DVAYS-DNC 119
DB 1457 -----NLCEMGQC-----LNVPGAYRCCECMG---FTPADSRSCQDIDEC 1494
QY 120 DV---IYVPGTDCN-----EEGYEL-----WTTDYDNI---PANCLN 150
DB 1495 SFONICVSGTCNNLPGMFHCICDDGELDRGTGNCYDIDECADPINCYN 1543

```

RESULT 11

```

ID XYNL_RUMFL STANDARD; PRT; 954 AA.
AC P29126;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bifunctional endo-1,4-beta-xylanase XYL4 precursor (EC 3.2.1.8).
GN XYNL.
OS Ruminococcus flavefaciens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Lachnospiraceae;
OC Ruminococcus.
OX NCBI_TaxID=1265;
RN [11]
RP SEQUENCE FROM N.A.
RC STRAIN=17;
RX MEDLINE=92261318; PubMed=1584021;
RA Zhang J.-X., Flint H.J.;

```

"A bifunctional xylanase encoded by the xynA gene of the rumen cellulolytic bacterium Ruminococcus flavefaciens 17 comprises two dissimilar domains linked by an asparagine/glutamine-rich sequence.";  
 Mol. Microbiol. 6:1013-1023(1992).  
 CC -1- FUNCTION: XYLANASE DOMAIN 1 RELEASES MORE XYL0-OLIGOSACCHARIDES  
 CC AND DOMAIN 2 MORE XYL0SE.  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-xylosidic  
 CC linkages in xylans.  
 CC -1- PATHWAY: xylan degradation.  
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY  
 CC G (FAMILY 11 OF GLYCOSYL HYDROLASES).  
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY  
 CC F (FAMILY 10 OF GLYCOSYL HYDROLASES).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

```

DR EMBL: Z11127; CAAT7476.1;
DR PIR: S20907; S20907.
DR HSSP: P48793; 1XND.
DR InterPro: IPR001000; Glyco_hydro_10.
DR InterPro: IPR001137; Glyco_hydro_11.
DR Pfam: PF00031; Glyco_hydro_10; 1.
DR Pfam: PF00457; Glyco_hydro_11; 1.
DR PRINTS: PR00134; GLHTRILASE10.
DR PRINTS: PR00911; GLHYDRLASE11.
DR SMART: SM00633; Glyco_10; 1.
DR PROSITE: PS00591; GLYCOSYL HYDROL_F10; 1.
DR PROSITE: PS00776; GLYCOSYL HYDROL_F11; 1.
DR PROSITE: PS00777; GLYCOSYL HYDROL_F12; 1.
KM xylan degradation; Hydrolase; Glycosidase; Multifunctional enzyme;
KW Repeat; Signal.
FT SIGNAL 1 27
FT CHAIN 28 954 OR 28, OR 29 (POTENTIAL).
FT DOMAIN 28 244 BIFUNCTIONAL ENDO-1,4-BETA-XYLANASE XYL4.
FT DOMAIN 245 622 XYLANASE DOMAIN 1.
FT DOMAIN 623 954 ASN/GLN/TRP-RICH (LINKER).
FT ACT_SITE 122 122 XYLANASE DOMAIN 2.
FT ACT_SITE 223 223 NOCLEOPHILE (BY SIMILARITY).
FT ACT_SITE 774 774 PROTON DONOR (BY SIMILARITY).
FT ACT_SITE 884 884 NUCLEOPHILE (BY SIMILARITY).
SQ SEQUENCE 954 AA; 111362 MW; 1033561DAB52EBD CRC64;

```

Query Match 8.88; Score 82.5; DB 1; Length 954;  
 Best Local Similarity 24.38; Pred. No. 34;  
 Matches 43; Conservative 26; Mismatches 37; Indels 71; Gaps 13;

```

QY 3 PWM---ADEANGAH--ODAW-KSLKADVENYVYMKATYKNDPWGNDPTCYGVANDV 56
DB 722 PDWFFRENSQNGAYSKIDIMNORLESMTKNTFALKSKQYPNIDVSYD-VC----- 772
QY 57 NEDKSIQAELFMNADPTMOPATEKVTAVKMYGYNR--ENAFRTEDGOVFTDVYAY 114
DB 773 NE-----LEFNNGG--GMRCG--DNSMWVKIYGDSFVYNAFKYARQ-----Y 811
QY 115 SDDNCVYIVVPGTDCNEGYELWTTDYDNIIPANC---LKNFNEYANGRETRVFTSA 168
DB 812 A-----PAGCKLYLNDYNEY-IPAKTNDIYNMA 838

```

RESULT 12

```

ID RPA2_DROME STANDARD; PRT; 1129 AA.
AC P20028;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE DNA-directed RNA polymerase I 135 kDa polypeptide (EC 2.7.7.6)

```



DE (RNA polymerase I subunit 2).  
GN Rpl135.  
OS Drosophila melanogaster (fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=9015849; PubMed=2482932;  
RT Kletterman R., Stizler S., Selfarth W., Petersen G., Bantz E.K.F.;  
RT "Primary structure and functional aspects of the gene coding for the  
RT second-largest subunit of RNA polymerase III of Drosophila.";  
RL Mol. Gen. Genet. 219:373-380(1989).  
CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION  
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS  
CC SUBSTRATES. RNA POLYMERASE I IS ESSENTIALLY USED TO TRANSCRIBE  
CC RIBOSOMAL DNA UNITS.  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +  
CC (RNA)(N).  
CC -1- SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM 9 TO 14  
CC DIFFERENT POLYPEPTIDES. THIS SUBUNIT IS THE SECOND LARGEST  
CC COMPONENT OF RNA POLYMERASE I.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES ARE  
CC FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA  
CC PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE  
CC III FOR 5S AND TRNA GENES.  
CC -1- SIMILARITY: Belongs to the RNA polymerase beta chain family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: X17298; CA35185.1; -.  
DR PIR: J00354; J00354.  
DR FlyBase: FBgn003278; Rpl135.  
DR InterPro: IPR001572; RNA\_pol\_1.  
DR Pfam: PF04563; RNA\_pol\_RpD2\_1; 1.  
DR Pfam: PF04561; RNA\_pol\_RpD2\_2; 1.  
DR Pfam: PF04565; RNA\_pol\_RpD2\_3; 1.  
DR Pfam: PF04567; RNA\_pol\_RpD2\_5; 1.  
DR Pfam: PF00562; RNA\_pol\_RpD2\_6; 1.  
DR Pfam: PF04560; RNA\_pol\_RpD2\_7; 1.  
DR PROSITE: PS01166; RNA\_POL\_BETA.1.  
KW Transferrase; DNA-directed RNA polymerase; Transcription; zinc;  
KW Zinc-finger; Nuclear protein.  
FT ZN.FING 1061 1093  
SQ SEQUENCE 1129 AA; 128414 MW; E0A15FE1BCEFB8D7 CRC64;  
  
Query Match 8.7%; Score 81.5; DB 1; Length 1129;  
Best Local Similarity 25.0%; Pred. NO. 22;  
Matches 21; Conservative 15; Mismatches 37; Indels 11; Gaps 3;  
  
OY 11 NGAHQDAWKSLLK--DVENYVYKATYKNDPV-----W---GNDFTCVGMANDVND 59  
DB 158 HGEHDSWGICFIVRGNEKIVRMILMTRNHPICVKKSSWMDRGNFSDLMVLQTVARE 217  
OY 60 EKSIQAEFLFMNNADTNNOFATEK 83  
DB 218 ESSLSNVVHYLNGTAKFMESHVK 241  
  
RESULT 13  
CICA\_PAEPP STANDARD: PRT; 695 AA.  
ID CICA\_PAEPP STANDARD: PRT; 695 AA.  
AC P57092;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Parasporal crystal protein cry18Ca (Parasporal delta-endotoxin  
DE Cryxviitic(a)) (Crystalline parasporal protoxin) (78 kDa crystal  
DE protein).  
GN CRY18CA OR CRYXVIITIC(A).  
OS Paenibacillus popilliae (Bacillus popilliae).  
OC Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus.  
OX NCBI\_TaxID=78057;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=ATCC 14706;  
RC Patel R., Yousten A.A., Rippere K.;  
RT "Detection of two new cry genes in Paenibacillus popilliae.";  
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
RL  
CC -1- FUNCTION: BINDS TO THE BRUSH BORDER MEMBRANE VESICLES OF SCARAB  
CC LARVAE AND DAMAGES THE GUT WALL, SOMEHOW TO ALLOW THE VEGETATIVE  
CC CELLS OF P. POPILLIAE TO ENTER THE HEMOLYMPH (BY SIMILARITY).  
CC -1- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING  
CC SPOROGATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART  
CC OF THE SPORE COAT (BY SIMILARITY).  
CC  
CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AF169251; AAF89668.1; -.  
DR InterPro: IPR005638; endotoxin\_C.  
DR InterPro: IPR005639; endotoxin\_N.  
DR Pfam: PF03944; endotoxin\_C\_1.  
DR Pfam: PF03945; endotoxin\_N\_1.  
KW Toxin; Sporulation.  
SQ SEQUENCE 695 AA; 78259 MW; 406AC9154D73E070 CRC64;  
  
Query Match 8.5%; Score 80; DB 1; Length 695;  
Best Local Similarity 24.5%; Pred. NO. 16;  
Matches 45; Conservative 13; Mismatches 68; Indels 58; Gaps 8;  
  
OY 18 WKSLEADVENYVYKATYKNDPVWNGDFTCVGMANDVNDKSIQAEFLFMNNADTNM 77  
DB 306 WSLK--YVNLVYVTSANLYN-----IGDNKNVNGEVSISWPFNFYIQKRS 351  
OY 78 QFATEKYAVKMYGYNRENAFRYETEDGQVFTDIAYSDNCDVIYYPGTDGNEGYELW 137  
DB 352 NYVLGVSQYAM-RMSYTNPE-----FGEYIDHLVNYTASVIGVNGPQIGQOLS 401  
OY 138 TT-----DYNNIPANC-----LNKFNEY-----AVGRERR-D 163  
DB 402 TTELDQLVQOQARADIPVDFTQIPINCLRNPLEVPYATRFNELTSLGTAGVGGEYRSD 461  
OY 164 VFTS 167  
DB 462 VFIS 465  
  
RESULT 14  
FBN2\_MOUSE STANDARD: PRT; 2907 AA.  
ID FBN2\_MOUSE STANDARD: PRT; 2907 AA.  
AC O61555; O63957;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Fibrillin 2 precursor.  
GN FBN2 OR FBN-2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]



FT	DISULFID	975	988	BY SIMILARITY.
FT	DISULFID	1070	1082	BY SIMILARITY.
FT	DISULFID	1077	1091	BY SIMILARITY.
FT	DISULFID	1093	1106	BY SIMILARITY.
FT	DISULFID	1112	1124	BY SIMILARITY.
FT	DISULFID	1119	1133	BY SIMILARITY.
FT	DISULFID	1135	1149	BY SIMILARITY.
FT	DISULFID	1155	1167	BY SIMILARITY.
FT	DISULFID	1162	1176	BY SIMILARITY.
FT	DISULFID	1178	1191	BY SIMILARITY.
FT	DISULFID	1197	1209	BY SIMILARITY.
FT	DISULFID	1204	1218	BY SIMILARITY.
FT	DISULFID	1220	1233	BY SIMILARITY.
FT	DISULFID	1239	1250	BY SIMILARITY.
FT	DISULFID	1246	1259	BY SIMILARITY.
FT	DISULFID	1261	1274	BY SIMILARITY.
FT	DISULFID	1280	1292	BY SIMILARITY.
FT	DISULFID	1287	1301	BY SIMILARITY.
FT	DISULFID	1303	1316	BY SIMILARITY.
FT	DISULFID	1322	1334	BY SIMILARITY.
FT	DISULFID	1329	1343	BY SIMILARITY.
FT	DISULFID	1345	1358	BY SIMILARITY.
FT	DISULFID	1364	1377	BY SIMILARITY.
FT	DISULFID	1371	1386	BY SIMILARITY.
FT	DISULFID	1388	1399	BY SIMILARITY.
FT	DISULFID	1405	1418	BY SIMILARITY.
FT	DISULFID	1412	1427	BY SIMILARITY.
FT	DISULFID	1429	1440	BY SIMILARITY.
FT	DISULFID	1446	1458	BY SIMILARITY.
FT	DISULFID	1453	1467	BY SIMILARITY.
FT	DISULFID	1469	1482	BY SIMILARITY.
FT	DISULFID	1488	1499	BY SIMILARITY.
FT	DISULFID	1494	1508	BY SIMILARITY.
FT	DISULFID	1510	1523	BY SIMILARITY.
FT	DISULFID	1529	1540	BY SIMILARITY.
FT	DISULFID	1535	1549	BY SIMILARITY.
FT	DISULFID	1551	1564	BY SIMILARITY.
FT	DISULFID	1647	1659	BY SIMILARITY.
FT	DISULFID	1654	1668	BY SIMILARITY.
FT	DISULFID	1670	1683	BY SIMILARITY.
FT	DISULFID	1689	1701	BY SIMILARITY.
FT	DISULFID	1696	1710	BY SIMILARITY.

	Query Match	8.5%	Score 79.5	DB 1	Length 2907	
	Match Similarity	29.0%	Pred. No. 98			
Batches	49	Conservative	15	Mismatches	46	Indels 59; Gaps 13.
OY	7 DEANGAHQDAMKSLKADVENVYVMVKATYK---- <td>63</td> <td></td> <td></td> <td></td> <td></td>	63				
	: : :   : :   :   :   :   :					
Dd	1403 DEACNGHQ---CSINACVN-----TBSYICACSSEGTGGFTC-----SDVDECAEN-	144.9				
	: : :   : :   :   :   :   :					
OY	64 QAEFLFMNMADTNMQFATEKTAVKMVGYNRENAFRVETEDGOVFT--DYIAYS-DNC	119				
	: : :   : :   :   :   :   :					
Dd	1450 -----TNLCENGQC-----LNVPGAYRCCEMG--FTPASDRSCDDIEDC	148				
	: : :   : :   :   :   :   :					
OY	120 DV--IYPGTDGN-----EEGYEL-----TTDYDN-I-PANCLN	150				
	: :                   :					
Dd	1489 SFONICVFCTCNNLPGMFHCI CDGDGYELDPRTGCNTDIDECADPINCYN	1537				
	: :                   :					
RESULT 15						
ITB4_HUMAN	STANDARD;	PRF:	1822 AA.			
ID ITB4_HUMAN	PI16144; O14690; O14691; O15339; O15340; O15341; Q9UIQ4;					
DT	01-APR-1990 (Rel. 14, Created)					
DT	16-OCT-2001 (Rel. 40, Last sequence update)					
DT	15-SEP-2003 (Rel. 42, Last annotation update)					
DE	Integrin beta-4 precursor (GP150) (CD104 antigen).					
GN	ITGB4					
OS	Homo sapiens (Human).					
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
CC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
OX	NCBI_Taxid:9606;					

RP [1]  
 RP SEQUENCE FROM N.A. (ISOFORM BETA-4A).  
 RX MEDLINE-90183973; PubMed-2311577;  
 RA Suzuki S., Naitoh Y.;  
 RT "Amino acid sequence of a novel integrin beta 4 subunit and primary  
 RT expression of the mRNA in epithelial cells.";  
 RL EMBO J. 9:757-763(1990).  
 RN [2]  
 RN SEQUENCE FROM N.A. (ISOFORM BETA-4B).  
 RX MEDLINE-90183974; PubMed-2311578;  
 RA Hogervorst F., Kulkman I., von Dem Borne A.E.G.K., Sonnenberg A.;  
 RT "Cloning and sequence analysis of beta-4 CDNA: an integrin subunit  
 RT that contains a unique 118 kd cytoplasmic domain.";  
 RL EMBO J. 9:765-770(1990).  
 RN [3]  
 RN SEQUENCE FROM N.A. (ISOFORM BETA-4C).  
 RP TISSUE-Pancreas;  
 RC MEDLINE-91009492; PubMed-1976638;  
 RX Tamura R.N., Kozzo C., Starr L., Chambers J., Reichardt L.F.,  
 RA Cooper H.M., Quaranta V.;  
 RT "Epithelial integrin alpha 6 beta 4: complete primary structure of  
 RT alpha 6 and variant forms of beta 4.";  
 RL J. Cell Biol. 111:1593-1604(1990).  
 RN [4]  
 RN SEQUENCE FROM N.A. (ISOFORMS BETA-4A; BETA-4B AND BETA-4C).  
 RP MEDLINE-97338298; PubMed-9194858;  
 RX Kulkinnen L., Kuriz K.-S., Xu Y., Bruckner-Tuderman L., Uitto J.;  
 RT "Genomic organization of the integrin beta 4 gene (ITGB4): a  
 RT homozygous splice-site mutation in a patient with junctional  
 RT epidermolysis bullosa associated with pyloric atresia.";  
 RL Lab. Invest. 76:823-833(1997).  
 RN [5]  
 RN SEQUENCE FROM N.A. (ISOFORMS BETA-4A; BETA-4B AND BETA-4C).  
 RP TISSUE-Lung;  
 RC MEDLINE-97311186; PubMed-9166594;  
 RX Tacovacci S., Gagnoux-Palacios L., Zamburino G., Meneguzzi G.,  
 RA D'Alessio M.;  
 RT "Genomic organization of the human integrin beta 4 gene.";  
 RL Mamm. Genome 8:448-450(1997).  
 RN [6]  
 RN REVISIONS.  
 RP D'Alessio M.;  
 RL submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RN SEQUENCE FROM N.A. (ISOFORM BETA-4E).  
 RX MEDLINE-97350870; PubMed-9207246;  
 RA van Leusden M.R., Kulkman I., Sonnenberg A.;  
 RT "The unique cytoplasmic domain of the human integrin variant beta4E is  
 RT produced by partial retention of intronic sequences.";  
 RL Biochem. Biophys. Res. Commun. 235:826-830(1997).  
 RN [8]  
 RN SEQUENCE OF 28-46.  
 RP MEDLINE-89251596; PubMed-2542022;  
 RA Kaij1 S., Tamura R.N., Quaranta V.;  
 RT "A novel integrin (alpha E beta 4) from human epithelial cells  
 RT suggests a fourth family of integrin adhesion receptors.";  
 RL EMBO J. 8:673-680(1989).  
 RN [9]  
 RN ALTERNATIVE SPLICING (ISOFORM BETA-4D).  
 RX MEDLINE-95073005; PubMed-7962032;  
 RA Clarke A.S., Lotz M.M., Mercurio A.M.;  
 RT "A novel structural variant of the human beta 4 integrin cDNA.";  
 RL Cell Adhes. Commun. 2:1-6(1994).  
 RN [10]  
 RN VARIANTS EB-PA TYR-61, CYS-252, ARG-562 AND TRP-1281.  
 RX MEDLINE-99011257; PubMed-9792864;  
 RA Gazzon M., Brown T., Smith L., Carter W.G., Uitto J.;  
 RT "Novel ITGB4 mutations in lethal and nonlethal variants of  
 RT epidermolysis bullosa with pyloric atresia: missense versus  
 RT nonsense.";  
 RL Am. J. Hum. Genet. 63:1376-1387(1998).  
 RN [11]

RP VARIANT EB-PA GUY-245.  
 RX MEDLINE-98082996; PubMed-9422533;  
 RA Pulkkinen L., Kim D.U., Uitto J.;  
 RT "Epidermolysis bullosa with pyloric atresia: novel mutations in the  
 RT beta-4 integrin gene (ITGB4)."  
 RL Am. J. Pathol. 152:157-166(1998).  
 RN [12]  
 RP VARIANT EB-PA PRO-156.  
 RX MEDLINE-98206430; PubMed-9546354;  
 RA Pulkkinen L., Buckner-Tuderman L., August C., Uitto J.;  
 RT "Compound heterozygosity for missense (L156P) and nonsense (R554X)  
 RT mutations in the beta-4 integrin gene (ITGB4) underlies mild,  
 RT nonlethal phenotype of epidermolysis bullosa with pyloric atresia."  
 RL Am. J. Pathol. 152:933-941(1998).  
 RN [13]  
 RP VARIANTS EB-PA ARG-38.  
 RX MEDLINE-99111354; PubMed-9892956;  
 RA Mellierio J.E., Pulkkinen L., McMillan J.R., Lake B.D., Horn H.M.,  
 RA Tidman M.J., Harper J.I., McGrath J.A., Uitto J., Eady R.A.J.;  
 RT "Pyloric atresia-junctional epidermolysis bullosa syndrome: mutations  
 RT in the integrin beta4 gene (ITGB4) in two unrelated patients with  
 RT mild disease."  
 RL Br. J. Dermatol. 139:862-871(1998).  
 RN [14]  
 RP VARIANT EB-PA TRP-1281.  
 RX MEDLINE-20334107; PubMed-10873890;  
 RA Kambham N., Tanji N., Seigle R.L., Markowitz G.S., Pulkkinen L.,  
 RA Uitto J., D'Agati V.D.;  
 RT "Congenital focal segmental glomerulosclerosis associated with beta4  
 RT integrin mutation and epidermolysis bullosa."  
 RL Am. J. Kidney Dis. 36:190-196(2000).  
 RN [15]  
 RP VARIANT GABER ASP-931.  
 RX MEDLINE-20235363; PubMed-10792571;  
 RA Inoue M., Tama K., Shimizu H., Owaribe K., Nakama T., Hashimoto T.,  
 RA McGrath J.A.;  
 RT "A homozygous missense mutation in the cytoplasmic tail of beta4  
 RT integrin, G931D, that disrupts hemidesmosome assembly and underlies  
 RT non-herlitz junctional epidermolysis bullosa without pyloric  
 RT atresia."  
 RL J. Invest. Dermatol. 114:1061-1064(2000).  
 RN [16]  
 RP VARIANTS EB-PA.  
 RX MEDLINE-21149723; PubMed-11251584;  
 RA Ashton G.H.S., Sorelli P., Mellierio J.E., Keane F.M., Eady R.A.J.,  
 RA McGrath J.A.;  
 RT "Alpha 6 beta 4 integrin abnormalities in junctional epidermolysis  
 RT bullosa with pyloric atresia."  
 RL Br. J. Dermatol. 144:408-414(2001).  
 RN [17]  
 RP VARIANTS HIS-98 AND LEU-844.  
 RX MEDLINE-21183353; PubMed-11289717;  
 RA Hirano A., Nagai H., Harada H., Terada Y., Haga S., Kajiwara T.,  
 RA Emi M.;  
 RT "Nine novel single-nucleotide polymorphisms in the integrin beta4  
 RT (ITGB4) gene in the Japanese population."  
 RL J. Hum. Genet. 46:35-37(2001).  
 RN [18]  
 RP VARIANTS EB-PA TYR-131; CYS-252; ASP-273; CYS-283; ASP-325; PRO-336  
 RP AND HIS-1225, AND VARIANT GLN-1216.  
 RX MEDLINE-21226999; PubMed-11328943;  
 RA Nakano A., Pulkkinen L., Murrell D., Ritco J., Lucky A.W., Garzon M.,  
 RA Stevens C.A., Robertson S., Pfendner E., Uitto J.;  
 RT "Epidermolysis bullosa with congenital pyloric atresia: novel  
 RT mutations in the beta 4 integrin gene (ITGB4) and genotype/phenotype  
 RT correlations."  
 RL Pediatr. Res. 49:618-626(2001).  
 CC -1- FUNCTION: INTEGRIN ALPHA-6/BETA-4 IS A RECEPTOR FOR LAMININ. IT  
 CC PLAYS A CRITICAL STRUCTURAL ROLE IN THE HEMIDESMOSOME OF  
 CC EPITHELIAL CELLS.  
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-4  
 CC ASSOCIATES WITH ALPHA-6.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named Isoforms-5;  
 CC Name-Beta-4C;  
 CC IsoId-P16144-1; Sequence-Displayed;  
 CC Name-Beta-4A;  
 CC IsoId-P16144-2; Sequence-VSP\_002749;  
 CC Name-Beta-4B;  
 CC IsoId-P16144-3; Sequence-VSP\_002749, VSP\_002750;  
 CC Name-Beta-4D;  
 CC IsoId-P16144-4; Sequence-VSP\_002749, VSP\_002751;  
 CC Name-Beta-4E;  
 CC IsoId-P16144-5; Sequence-VSP\_002747, VSP\_002748;  
 CC TISSUE SPECIFICITY: INTEGRIN ALPHA-6/BETA-4 IS PREDOMINANTLY  
 CC EXPRESSED BY EPITHELIA. ISOFORM BETA-4E IS ALSO EXPRESSED IN COLON  
 CC AND PLACENTA. ISOFORM BETA-4E IS ALSO EXPRESSED IN EPIDERMIS,  
 CC LUNG, DUODENUM, HEART, SPLEEN AND STOMACH.  
 CC -1- DOMAIN: THE FIBRONECTIN TYPE III-LIKE DOMAINS BIND BPAG1 AND  
 CC PLECTIN AND PROBABLY ALSO RECRUIT BP230.  
 CC -1- DISEASE: DEFECTS IN ITGB4 ARE A CAUSE OF EPIDERMOLYSIS BULLOSA  
 CC LETALIS WITH PYLORIC ATRESIA (EB-PA), ALSO KNOWN AS JUNCTIONAL  
 CC EPIDERMOLYSIS BULLOSA WITH PYLORIC ATRESIA (PA-JEB) OR APIASIA  
 CC CUTIS CONGENITA WITH GASTROINTESTINAL ATRESIA. IT IS CHARACTERIZED  
 CC BY MUCCUTANEOUS FRAGILITY AND GASTROINTESTINAL ATRESIA, WHICH  
 CC MOST COMMONLY AFFECTS THE PYLORUS.  
 CC -1- DISEASE: DEFECTS IN ITGB4 ARE A CAUSE OF GENERALIZED ATROPHIC  
 CC BENIGN EPIDERMOLYSIS BULLOSA (GABE). THIS NONLETHAL FORM OF  
 CC JUNCTIONAL EPIDERMOLYSIS BULLOSA IS CHARACTERIZED BY LIFE-LONG  
 CC BLISTERING OF THE SKIN, ASSOCIATED WITH HAIR AND TOOTH  
 CC ABNORMALITIES.  
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.  
 CC -1- SIMILARITY: Contains 2 vWFA-like domains.  
 CC -1- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN IN POSITIONS 1414  
 CC TO 1429 DUE TO A FRAMESHIFT.  
 CC -1- DATABASE: NAME-PROV; NOTE-CD guide CD104 entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd104.htm".  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

Query Match 8.48; Score 78.5; DB 1; Length 1822;  
 Best Local Similarity 23.18; Pred. No. 69;  
 Matches 28; Conservative 20; Mismatches 50; Indels 23; Gaps 5;

OY 42 WGN-----DPTCYGVANVDNEDEKSIQAEFLFMN-----NADTMWQFAT-----EKV 84  
 Db 194 WPNSDPPFSFKVNI SLTEDYDEFRNKLQGERISGNLDAPEGGDAILQTVCTRIDGWRP 253  
 OY 85 TAVKMYGNRENAFRYETEDGQVFTDVIAVSDNCDVIVYVPGTGNEGEELMTTDDYDNI 144  
 Db 254 DSHLLLVFSPESAFHREADGANYLACIMSNDRCHL-----DPTGYTYGR--TQYPSV 307  
 OY 145 P 145  
 Db 308 P 308

Search completed: August 1, 2003, 12:15:51  
 Job time : 13.4229 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: August 1, 2003, 12:15:04 ; Search time 12.7029 Seconds  
(without alignments)  
1294.578 million cell updates/sec

Title: US-10-087-195-7

Perfect score: 936

Sequence: 1 NQPMADDEANGAHQDAWKS.....FNEYAVGRETROVETSACLE 171

Scoring table: BL0SUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	105	11.2	1592	2 A38175	glucosyltransferase
2	91	9.7	450	2 H82891	hypothetical prote
3	88	9.4	606	2 H97012	hypothetical prote
4	86.5	9.2	254	2 T18282	hypothetical prote
5	85.5	9.1	1237	2 AC1583	internalin protein
6	85.5	9.1	1612	2 AB1347	probable peptidogl
7	84.5	9.0	858	1 A42239	adenylate cyclase
8	83.5	8.9	2918	2 A54105	fibrillin-2 precur
9	83	8.9	601	2 S57962	casp protein - Clo
10	82.5	8.8	954	1 S20907	endo-1,4-beta-xyla
11	81.5	8.7	1129	2 J00354	DNA-directed transc
12	80	8.5	166	2 T51637	myb-related transc
13	80	8.5	249	2 E84717	probable MYB fami
14	79.5	8.5	362	2 A12159	hypothetical prote
15	79.5	8.5	1377	2 D90538	hypothetical prote
16	79.5	8.5	2907	2 A57278	fibrillin-2 precu
17	79	8.4	327	2 D89931	branched-chain alp
18	79	8.4	1872	2 T30888	vitellogenin - Ath
19	78.5	8.4	531	2 A84471	En/Spm-like transp
20	78.5	8.4	785	2 H90547	lipoprotein (impor
21	78.5	8.4	964	2 JC5845	integrin beta-4 pr
22	78.5	8.4	1875	2 A36429	L-lactate dehydrog
23	78	8.3	317	2 A11206	hypothetical prote
24	78	8.3	338	2 C86775	hypothetical prote
25	78	8.3	363	2 T20745	hypothetical prote
26	78	8.3	446	2 G84262	prophage p13 prote
27	78	8.3	595	2 B86788	hypothetical prote
28	77.5	8.3	421	1 A38168	glutamate dehydrog
29	77.5	8.3	512	2 S37279	hexon protein - hu

30	77.5	8.3	1026	2 A49750	beta-galactosidase
31	77.5	8.3	1659	2 H97926	hypothetical prote
32	77.5	8.3	1659	2 G95057	endo-beta-N-acetyl
33	77	8.2	150	2 A59103	hypothetical prote
34	77	8.2	236	2 E89769	hypothetical prote
35	77	8.2	335	2 G64385	hypothetical prote
36	77	8.2	366	2 F97343	uncharacterized co
37	77	8.2	393	2 S06256	hypothetical prote
38	77	8.2	2893	2 A64556	toxin-like outer m
39	76.5	8.2	669	2 E71610	hypothetical prote
40	76.5	8.2	1219	2 H84464	probable helicase
41	76	8.1	153	2 T06182	reverse transcript
42	76	8.1	368	1 Q0BEHG	early nuclear anti
43	76	8.1	403	2 T09322	DNA polymerase pro
44	76	8.1	466	2 G72603	nitrate reductase
45	76	8.1	515	2 H72455	hypothetical prote

## ALIGNMENTS

## RESULT 1

A38175 glucosyltransferase precursor - Streptococcus sobrinus

C:Species: Streptococcus sobrinus

C>Date: 28-Aug-1992 #sequence\_revision 28-Aug-1992 #text\_change 15-Oct-1999

C:Accession: A38175

R:Abu, H.; Msumura, T.; Kodama, T.; Ohta, H.; Fukui, K.; Kato, K.; Kagawa, H.

J. Bacteriol. 173, 989-996, 1991

A:Title: Peptide sequences for sucrose splitting and glucan binding with in Streptococ

A:Reference number: A38175; MUID:91123227; PMID:1704006

A:Accession: A38175

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1592 <ABO>

A:Cross-references: GB:D90213; NID:g217032; PIDN:BA14241.1; PID:d1014946; PID:g21703

C:Superfamily: cpl repeat homology

F:1093-1112/Domain: cpl repeat homology <CP1>

F:1222-1241/Domain: cpl repeat homology <CP2>

F:1287-1306/Domain: cpl repeat homology <CP3>

F:1350-1351/Domain: cpl repeat homology <CP4>

F:1352-1371/Domain: cpl repeat homology <CP5>

F:1402-1420/Domain: cpl repeat homology <CP6>

F:1465-1484/Domain: cpl repeat homology <CP7>

F:1513-1532/Domain: cpl repeat homology <CP8>

Query Match 11.2% Score 105; DB 2; Length 1592;

Best Local Similarity 25.4% Pred. No. 0.52;

Matches 51; Conservative 19; Mismatches 55; Indels 76; Gaps 12;

Oy	2	QPMW-----ADEANGAHQ-DAMKSLKADVENYVM-----KATYK-N 38
Db	335	QPMWNGSEKRYDHLONGALLFDNGTDLPTDQSNRYLNRPTNQTGSLDSRFTYNNP 394
Oy	39	DPVWNGDFTGCVANVDNEDEKSIQAEFL-----FNNNADTNQFATEKVTA 86
Db	395	DPVWNGDFTGCVANVDNEDEKSIQAEFL-----FNNNADTNQFATEKVTA 86
Oy	87	VKMTGVRENARFRETEDGVFTDVIASDDNCVIVYGTGNGECC-----YELWT- 138
Db	442	-----DSIRVDAED-NVDADQLOIISDYLAAY--GIDKNNKNNNNHVSIVEAMSD 489
Oy	139	-----TDVNIIPANCLNKF 152
Db	490	NDRPYLHDDGDL-MNMDNKF 509

## RESULT 2

H82891 hypothetical protein U0441 [imported] - Ureaplasma urealyticum

C:Species: Ureaplasma urealyticum

C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000

C:Accession: H82891

R;Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
Submitted to Genbank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mlt  
A:Reference number: A82870  
A:Accession: H82891  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-450 <GLA>  
A:Cross-references: GB:AE002140; GB:AF222894; NID:96899420; PIDN:AAF30853.1; GSPDB:GN001  
A:Experimental source: serovar 3; biovar 1  
C:Genetics:  
A:Gene: UM441  
A:Genetic code: SGC3

Query Match 9.7%; Score 91; DB 2; Length 450;  
Best Local Similarity 23.4%; Pred. No. 1.9;  
Matches 37; Conservative 23; Mismatches 60; Indels 38; Gaps 8;

DB 7 DEANAGHDAKSLKADVENYVYKATYKNDPVGNDFTCVG---VANDVNEDEKS 62  
19 DENSESHIRKIMELMVLQNDVKIINHSYGWSP---NDLNLGIFYNFNGKEIN-DDKD 74  
QY 63 IQAEFL-----FMNADTNMOPATEKVTAVKMY-----GYNENAFRYE 101  
DB 75 FKRELKRYIDTLKNDIVANTINKYFKKATYFISYFPIITNIFSAGNSYND--IRTR 132  
QY 102 TEDGV---FTDVIAVS---DDNCVIVYVPGTDGNEE 132  
DB 133 NKNGHILEINWFDISKYSFMDREFNONSIVYSTINKNE 170

RESULT 3  
H97012  
hypothetical protein CAC0915 [imported] - Clostridium acetobutylicum  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
C:Accession: H97012  
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld  
A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: H97012  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-606 <RUR>  
A:Cross-references: GB:AE001437; PIDN:AAK78891.1; PID:915023815; GSPDB:GN00168  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC0915

Query Match 9.4%; Score 88; DB 2; Length 606;  
Best Local Similarity 25.5%; Pred. No. 5;  
Matches 27; Conservative 15; Mismatches 34; Indels 30; Gaps 5;

QY 55 DVNEDKSKIOAEFLFMNADTNMOPATEKVTAVKMGYNENAFRYETEDGVFTDVIAV 114  
DB 40 DTNAERQAI-SPYIY---GTNODFSNAKYTA-RRIGGRKSTGYNNENNDNSNA----- 86  
QY 115 SDNDCVIVYVPGTDGNEEGVELMTDYD-----NIPANCLNKENE 154  
DB 87 -----GTDMKNESDNTWLTLYDPKEXYNEPASVYTAFFHD 121

RESULT 4  
T18282  
hypothetical protein G6 - slime mold (Dictyostelium discoideum) plasmid  
C:Species: Dictyostelium discoideum  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 01-Dec-2000  
C:Accession: T18282; S28718  
R;Rieken Jr., W.K.; Gonzales, C.M.; Gonzales, S.T.; Pilkington, K.J.; Kiyosawa, H.; Hugh  
Genetics 148, 1117-1125, 1998  
A:Title: Dictyostelium discoideum nuclear plasmid Ddp5 is a chimera related to the Ddp1

A:Reference number: Z14684; MUID:98198836; PMID:9539429  
A:Accession: T18282  
A>Status: preliminary; translated from GB/EMBL/DDJ  
A:Molecule type: DNA  
A:Residues: 1-254 <RIE>  
A:Cross-references: EMBL:000796; NID:92702254; PID:92702257; PIDN:AAC18633.1  
A:Experimental source: plasmid Ddp5  
R;Gurniak, C.B.; Bang, A.G.; Noegel, A.A.  
Curr. Genet. 17, 321-325, 1990  
A:Title: Transcript and sequence analysis of a 5.1 kb contiguous fragment of Dictyost  
A:Reference number: S28718; MUID:90254839; PMID:2340592  
A:Accession: S28718  
A:Molecule type: DNA  
A:Residues: 37-254 <GUR>  
A:Cross-references: EMBL:X53237; NID:97280; PIDN:CAA37327.1; PID:97281  
A:Experimental source: Plasmid Ddp1  
C:Genetics:  
A:Gene: d-5  
A:Genome: plasmid

Query Match 9.2%; Score 86.5; DB 2; Length 254;  
Best Local Similarity 23.5%; Pred. No. 2.3;  
Matches 31; Conservative 12; Mismatches 48; Indels 41; Gaps 5;

QY 26 ENYVYKATYKNDPVGNDFTCVGVANDVNEDEKSIOAEFLFMNADTNMOPATEKYT 85  
DB 155 ENYVHTNITIOEK---SNRFCTDPLCHYCKNE-----NIQNLDFKTKCT 198  
QY 86 AVKMGYNRNFARFETEDGVFTDVIAVSDNDCVIVYVPGTDGNEEGVELMTDYDNIIP 145  
DB 199 P-----KYGASDSEFLS-----TTYNPKLDGSSNGKSVTOEKNIS 235

QY 146 ANCLNKNEXAV 157  
DB 236 NNL-KINITILI 245

RESULT 5  
AC1583  
interalin protein (LPXYG motif) homolog lln1204 [imported] - Listeria innocua (strai  
C:Species: Listeria innocua  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AC1583  
R;Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec  
J. Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl,  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkut, G.; Madueno, E.; Maltournam, A.;  
Ok, C.; Schluter, T.; Simoes, N.; Tilleret, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AC1583  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1237 <GLA>  
A:Cross-references: GB:AL592022; PIDN:CAC96435.1; PID:916413678; GSPDB:GN00178  
A:Experimental source: strain Clp11262  
C:Genetics:  
A:Gene: lln1204

Query Match 9.1%; Score 85.5; DB 2; Length 1237;  
Best Local Similarity 27.3%; Pred. No. 21;  
Matches 41; Conservative 18; Mismatches 70; Indels 21; Gaps 7;

QY 21 LKADVENYVYKKA-----TYKNDPVGNDFTCVGVANDVNEDEKSIOAEFL--FMNN 72  
DB 788 IKATPKREVIVYIDAAGANITVYKEDSCNKAENSLTGNVGEETSSSAKELGTTTLE 847  
QY 73 ADTNMQ--FATEKVTAVKMGYN---RENARFETEDGVFTDVIAVSDNDCVIVYVPG 126  
DB 848 MFSNMGESLSLEQVITYYSKNPVPKAKITQVYTDDEDMELAPFTTLS-GAVDEVYVA- 905  
QY 127 TDGNEEGVELMTDYDNIIPANCLNKNEXA 156

Db 906 TAKFTGTGELIET-----PSNAEGKFESENA 930

RESULT 6

AB1347

probable peptidoglycan bound protein (LPXTG motif) lmo2178 [imported] - Listeria monocytogenes

C:Species: Listeria monocytogenes

C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001

C:Accession: AB1347

R:Glaser, P.; Fraigneul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl, H.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maitournam, A.; Meck, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AB1347

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1612 <GLA>

A:Cross-references: GB:NC\_003210; PIDN:CAD00256.1; PID:g16411648; GSPDB:GN00177

A:Experimental source: strain EGD-e

A:Genetics:

A:Gene: lmo2178

Query Match 9.18; Score 85.5; DB 2; Length 1612;

Best Local Similarity 26.08; Pred. No. 29;

Matches 44; Conservative 22; Mismatches 68; Indels 35; Gaps 9;

Db 4 DMADEA--ANGAHODAMKSLKADYENYVYWKATYKNDPVWGNDFTCVGVANDVNEDEK 61

513 DMTSATSANONENFLTKNTNTNAYELITSKIDFSRKTKNET---DENGVASDAT 568

Oy 62 -SIAEFL-----FMNADTNMGRATEKVA---VKWGYNRENAFRYETEDGVF 108

Db 569 IAIPDLKKEAGTIDFENNT-----MTWKITANSRDKMGNINTEDEFSTGVKALSKSY 622

Oy 109 TDVAYSDDNCDVIYVPGTQNEGCELTMTDYDNIPAN-CLNKFNEVA 156

Db 623 T-VRAVYDNTNSVLTETGKD-----YTIDKVTAPAGFYIOLIGDYA 662

RESULT 7

A42239

adenylate cyclase (EC 4.6.1.1) germination stage - slime mold (Dictyostelium discoideum)

C:Species: Dictyostelium discoideum

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C:Accession: A42239

R:Plitt, G.S.; Milona, N.; Borleis, J.; Lin, K.C.; Reed, R.R.; Devreotes, P.N.

Cell 69, 305-315, 1992

A:Title: Structurally distinct and stage-specific adenylyl cyclase genes play different

A:Reference number: A42239; MUID:92233467; PMID:1348970

A:Accession: A42239

A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra

A:Molecule type: DNA

A:Residues: 1-858 <PIT>

A:Cross-references: GB:M87278

C:Superfamily: slime mold germination stage adenylyl cyclase; guanylate cyclase catalyt

C:Keywords: phosphorus-oxygen lyase

F:351-574/Domain: guanylate cyclase catalytic domain homology <GCC>

Query Match 9.0%; Score 84.5; DB 1; Length 858;

Best Local Similarity 23.28; Pred. No. 16;

Matches 39; Conservative 31; Mismatches 61; Indels 37; Gaps 9;

Db 5 MADEANAGAHODAMKSLKADYENY-----VTM-----VKATYKNDPVWG-NDFTCVGVA 53

43 YSEBENSINQMDLERSSQIITHNQNMAYLLSSIDLKALYYVNPFRDRDPNFV---- 98

Oy 54 NDVNEDEKSIAEFLF-----MNNADTNM---QFATEKTYAVKMGYVNE-NAFRYETEDG 105

Db 99 --LNTLTKNSEFOYLFIKKIKNNNDRCNCFEERKFKIDPFQIYSFDEMTNSIHVANKKS 156

Oy 106 QVFMDVATSDNCDVYVPGTQNEGCELTMTDYDN--IPANCLNK 151

Db 157 SYFPIIAFPDINKDII-----GLDINSTDYANETIKKSIKFNK 194

RESULT 8

A54105

fibrillin-2 precursor - human

C:Species: Homo sapiens (man)

C:Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 02-Aug-2002

C:Accession: A54105; S17063; S31101

R:Zhang, H.; Pfeleth, S.D.; Hu, W.; Davis, E.C.; Sangunetti, C.; Bonadio, J.; Mecha

J. Cell Biol. 124, 855-863, 1994

A:Title: Structure and expression of fibrillin-2, a novel microfibrillar component pr

A:Reference number: A54105; MUID:94165150; PMID:8120105

A:Accession: A54105

A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual

A:Molecule type: mRNA

A:Residues: 1-2918 <ZHA>

A:Cross-references: GB:U03272

R:Lee, B.; Godfrey, M.; Vitale, E.; Horl, H.; Mattel, M.G.; Sarfarazi, M.; Tsipouras,

Nature 352, 330-334, 1991

A:Title: Linkage of Marfan syndrome and a phenotypically related disorder to two diff

A:Reference number: S17062; MUID:91304567; PMID:1852206

A:Accession: S17063

A:Molecule type: mRNA

A:Residues: 752-1489, 1791, 'GS', 1794-1796, 'QLT', 1922-1923, 'LD', 1926, 'P', 1928 <LEE>

A:Cross-references: EMBL:X62009

R:Miliewicz, D.M.

submitted to the EMBL Data Library, December 1992

A:Accession: S31101

A:Reference number: S31101

A:Molecule type: mRNA

A:Residues: 752-1407, 'R', 1409-1489, 1791, 'GS', 1794-1796, 'QLT', 1922-1923, 'LD', 1926, 'P',

A:Cross-references: EMBL:X62009

C:Genetics:

A:Gene: GDB:FBN2

A:Cross-references: GDB:128122; OMIM:121050

A:Map position: 5q23-5q31

C:Superfamily: fibrillin 1; EGF homology

C:Keywords: extracellular protein

F:1-29/Domain: signal sequence #status predicted <SIG>

F:30-2918/Product: fibrillin-2 #status predicted <MNT>

F:1245-1280/Domain: EGF homology <EGF>

F:1970-2013/Domain: EGF homology <EGF>

Query Match 8.9%; Score 83.5; DB 2; Length 2918;

Best Local Similarity 29.0%; Pred. No. 95;

Matches 49; Conservative 16; Mismatches 45; Indels 59; Gaps 13;

Oy 7 DEANGAHODAMKSLKADYENYVYWKATYR---NDPVGNDFTCVGVANDVNEDEKSI 63

Db 1409 DECSNGTHQ---CSINACVN---TPESYRCACSEGTGGGFTC---SDVCECAENI 1456

Oy 64 QAEFLNMMADTNMGRATEKVAVKMGYVNEENAFRYETEDGVFT---DVIAYSD-DNC 119

Db 1457 -----NLCEKGC-----LNPAGVRCCEMG---FTPASDSRSCODIDEC 1494

Oy 120 DV--IYVPGTGN-----ERGYEL---WTTDYDNI--PANCLN 150

Db 1495 SFONICVSGTCNNLPGMFHCICDDGYELDRGNGCTDIDECADPINCVA 1543

RESULT 9

S57962

cspc protein - Clostridium acetobutylicum (fragment)

C:Species: Clostridium acetobutylicum

C:Date: 13-Jan-1996 #sequence\_revision 19-Apr-1996 #text\_change 23-Mar-2001

C:Accession: S57962

R:Sanchez-Beato, A.; Garcia, J.

submitted to the EMBL Data Library, July 1995

A:Description: Molecular characterization of a family of choline-binding proteins of *Cld*  
A:Reference number: S57714  
A:Accession: S57962  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-601 <S>  
A:Cross-references: EMBL:Z50033; NID:9899232; PIDN:CAA90347.1; PID:9899233  
C:Superfamily: cpl repeat homology  
F:533-552/Domain: cpl repeat homology <CP3>  
F:553-572/Domain: cpl repeat homology <CP4>

Query Match  
Best Local Similarity 27.0%; Pred. No. 14;  
Matches 31; Conservative 14; Mismatches 32; Indels 38; Gaps 6;

QY 57 NEDKSIQAEFLPMN-----NADTNMOPATEKVAVMGYNRE-----NAFRYET 102  
DB 324 DEQKNDKADFLNTGTGTINEDTKFVANGKLA-----YNTLMNKYKVAAYSLKT 378  
QY 103 EDGQVFTDVIAYSDNDV-----IYVPGTDGNEGYELWTTD-----YDN 143  
DB 379 KGGYYADEDESKEDKESVSDDKTSVOTDVGN-----LWRLDGYIKFEDN 428

## RESULT 10

S20907  
endo-1,4-beta-xylanase (EC 3.2.1.8) precursor, bifunctional - *Ruminococcus flavefaciens*  
N:Contains: endo-1,4-beta-xylanase (EC 3.2.1.8)  
C:Species: *Ruminococcus flavefaciens*  
C>Date: 22-Nov-1993 #sequence\_revision 22-Nov-1996 #text\_change 18-Jun-1999  
C:Accession: S20907; S18043  
R:Zhang, J.X.; Flint, H.J.  
Mol. Microbiol. 6, 1013-1023, 1992  
A:Title: A bifunctional xylanase encoded by the xyna gene of the rumen cellulytic bacterium.  
A:Reference number: S20907; MUID:92261318; PMID:1584021  
A:Accession: S20907  
A:Molecule type: DNA  
A:Residues: 1-954 <ZHA>  
A:Cross-references: EMBL:Z11127; NID:946161; PIDN:CAA77476.1; PID:9581505  
A:Experimental source: strain 17  
A:Genetics:  
A:Gene: xyna  
A:Start codon: TTG  
A:Function:  
C:Description: catalyzes the hydrolysis of 1,4-beta-xylosidic bonds in xylans  
A:Pathway: xylan degradation  
C:Superfamily: *Ruminococcus* bifunctional endo-1,4-beta-xylanase; endo-1,4-beta-xylanase  
C:Keywords: extracellular protein; glycosidase; hydrolase; multifunctional enzyme; polys  
F:1-27/Domain: signal sequence #status predicted <SIG>  
F:28-954/Product: bifunctional endo-1,4-beta-xylanase #status predicted <MAT>  
F:28-248/Domain: catalytic domain A #status predicted <CDA>  
F:40-236/Domain: endo-1,4-beta-xylanase homology <XYL>  
F:249-622/Domain: linker region B #status predicted <LRB>  
F:623-954/Domain: catalytic domain C #status predicted <CDC>  
F:655-953/Domain: streptomycin endo-1,4-beta-xylanase A homology <XYL>  
F:122,223/Active site: Glu #status predicted

Query Match  
Best Local Similarity 24.3%; Pred. No. 28;  
Matches 43; Conservative 26; Mismatches 37; Indels 71; Gaps 13;

QY 3 PDW---ADEAANGAH---QDAW-KSLKADVENYVYKATYKNDPVWGNDFTCVGVANDV 56  
DB 722 PDWFEFNSQNGAYVSKDIMNQRLESMIKNTFAKSOVPNDVSYD-VC----- 772  
QY 57 NEDKSIQAEFLPMNADTNMOPATEKVAVMGYNRE-----NAFRYETGQVFTDVIAY 114  
DB 773 NE-----LELNGG-GMRGA-DNSNMVKYIGDSEFVINAFAKARQ-----Y 811  
QY 115 SDNCDVIYVPGTDGNEGYELWTTDYNIPANC-----LNKFNAYAVGRETRDVFYSA 168  
DB 812 A-----PACCKLYLNDYNEY-IPAKTNDIYVNA 838

## RESULT 11

J00354  
DNA-directed RNA polymerase (EC 2.7.7.6) I second largest chain - fruit fly (*Drosophila*)  
C:Species: *Drosophila melanogaster*  
C>Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 21-Jul-2000  
C:Accession: J00354; S07025  
R:Kontemann, R.; Stitzler, S.; Selfarth, W.; Petersen, G.; Bautz, E.K.F.  
Mol. Gen. Genet. 219, 373-380, 1989  
A:Title: Primary structure and functional aspects of the gene coding for the second-1  
A:Reference number: J00354; MUID:90158499; PMID:2482932  
A:Accession: J00354  
A:Molecule type: DNA  
A:Residues: 1-1129 <KON>  
A:Cross-references: EMBL:X17298; NID:98473; PIDN:CAA35185.1; PID:98474  
A>Note: the authors translated the codon CAG for residue 202 as Glu, TCT for residue  
C:Genetics:  
A:Gene: DmRP135  
A:Cross-references: FlyBase:FBgn0003278  
A:Introns: 30/3; 530/3  
C:Superfamily: DNA-directed RNA polymerase 132K polypeptide  
C:Keywords: DNA binding; nucleotidyltransferase; nucleus; transcription; zinc finger

Query Match  
Best Local Similarity 25.0%; Pred. No. 42;  
Matches 21; Conservative 15; Mismatches 37; Indels 11; Gaps 3;

QY 11 NGAHQDWMKSLKA--DVENVYVYKATYKNDPV-----W--GNDFTCVGVANDVNE 59  
DB 158 HGEHSEMGIFVIRGNEKTVRLIMTRNHPICVRRSSWKRQNSDLGLVQTVRED 217  
QY 60 EKSIQAEFLPMNADTNMOPATEK 83  
DB 218 ESSLNSVHYHLNNGTAKMFMSHVK 241

## RESULT 12

T51637  
myb-related transcription factor MYB14 [imported] - *Arabidopsis thaliana* (fragment)  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
C>Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 17-Nov-2000  
C:Accession: T51637  
R:Kranz, H.D.; Denekamp, M.; Greco, R.; Jin, H.; Kranz, H.D.; Denekamp, M.; Greco, R.  
Plant J. 16, 263-276, 1998  
A:Title: Towards functional characterisation of the members of the R2R3-MYB gene from  
A:Reference number: Z14349; MUID:9839469; PMID:9839469  
A:Accession: T51637  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-166 <KRA>  
A:Cross-references: EMBL:AF062865; PIDN:ACG3587.1  
A:Experimental source: cultivar Columbia  
C:Genetics:  
A:Gene: MYB14  
C:Superfamily: *Arabidopsis* 28K leaf-specific myb-related protein; myb DNA-binding rep  
C:Keywords: transcription factor

Query Match  
Best Local Similarity 22.6%; Pred. No. 5;  
Matches 38; Conservative 27; Mismatches 65; Indels 38; Gaps 7;

QY 18 WKSILKA-----DVENVYVYV-KVATYKNDPVWGNDFTCVGVANDVNEDEKSIQAEFL 68  
DB 6 WSAIAKILGRDNDNEIKNWNHTLKRSLKLNLSNGDPTADVNGINTETEDKGSVYVD-- 63  
QY 69 FNNADTNMOPATEKVAVMGYNRENAFRYETEDGQVFTDVIAYSDN--CDVITYPG 126  
DB 64 -----TASIQOFSNITFDISNDKDDINSYE-----DISALIDDSFMSDVYSVN 110  
QY 127 TIDGNEGYELWTTDYNIPANC-----LNKFNAYAVGRETRDVFYTS 167



Db 111 SNRKEKIEDMEGLIDRNSKCSYSNSKLYNDMEFNF-----DVFTS 153

## RESULT 13

probable MYB family transcription factor [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 17-May-2002

C:Accession: E84717

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shee, T.P.; Baito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.; Euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: E84717

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-249 <STO>

A:Cross-references: GB:AE002093; NID:g4432813; PIDN:AAD20663.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g31180

A:Map position: 2

C:Superfamily: Arabidopsis 28k leaf-specific myb-related protein; myb DNA-binding repeat

Query Match 8.5%; Score 80; DB 2; Length 249;

Best Local Similarity 22.6%; Pred. No. 8.4;

Matches 38; Conservative 27; Mismatches 65; Indels 38; Gaps 7;

18 WKSILKA-----DVENVYVM-VKATYKNDPVGNDFTCVGVANDVEDKSIQAEFL 68

Db 89 WSAIAAKLPRTDNEINVMHTLKRSLKLNNGCPTKVCNGINETNEDKSGSIYD-- 146

Qy 69 FMNADNMGPATEKVTAVMYGYNRENAFRYETEDGVFTDVAISDDN--CGVIYVPG 126

Db 147 -----TASLDQFSSTTFDINSNDKDDINSTF-----DISALIDDSWSDVSYDN 193

Qy 127 TDGNEGEYELMTTDYDNIIPANC-----LNKFNVAVGRTRDVF 167

Db 194 SNRKEKIEDMEGLIDRNSKCSYSNSKLYNDMEFNF-----DVFTS 236

RESULT 14

A:Accession: A12159

hypothetical protein alr2832 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002

C:Accession: A12159

R:Kanehiko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi, N.; Shimo, S.; Sugimoto, M.; Takezawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: A12159

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-362 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA074531.1; PID:g17131926; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: alr2832

C:Superfamily: hypothetical protein alr1534

Query Match 8.5%; Score 79.5; DB 2; Length 362;

Best Local Similarity 23.6%; Pred. No. 15;

Matches 42; Conservative 26; Mismatches 69; Indels 41; Gaps 9;

23 ADVENVYVMKA---TYKNDPVGNDFTCVG---VMANDVEDKSIQAEFL---FMNN 72

Db 192 SDTDEYTLASARMSISTKRDIYEAFLNMLGWLPLISGDGPERRR-LQAKALDNKFLGH 250

Qy 73 ADTNMQ---FATER---VTAVKMGYNRENAFRYETEDGVFTDVAISDNCVYVPG 126

Db 251 VSDNRKRELFSSKASIIVALEDEYGL-----VPVEANASGRPVIAFGAGVLDITQING 303

Qy 127 TDG-----NEGEYELMTTDYDNIIPANC-LNKFNVAVGRTRDVF 169

Db 304 QTGVFFKRPQPELSOKALLESGEITW--DYENIRNNAVNMFSEVPFSKVERVITQIC 359

## RESULT 15

hypothetical protein MYPU\_2120 [imported] - Mycoplasma pulmonis (strain UAB CTIP)

C:Species: Mycoplasma pulmonis

C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001

C:Accession: D90538

R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Gallisson, F.; Moszer, Nucleic Acids Res. 29, 2145-2153, 2001

A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma p

A:Reference number: A9512; MUID:21267165; PMID:11353084

A:Accession: D90538

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1377 <KUR>

A:Cross-references: GB:AL445566; PID:g14089625; PIDN:CAC13385.1; GSPDB:GN00153

A:Experimental source: strain UAB CTIP

C:Genetics:

A:Gene: MYPU\_2120

A:Genetic code: SCC3

Query Match 8.5%; Score 79.5; DB 2; Length 1377;

Best Local Similarity 24.7%; Pred. No. 83;

Matches 42; Conservative 27; Mismatches 50; Indels 51; Gaps 11;

15 QDAWKSIAKADVENYVMKAT-----YKND-----PYMGND-----FTCVGMAN-D 55

Db 1116 QDAWKS-----EDKYVYDOKANQNTIASQYKLDQKVKIRPNKSDALKETFLGTTINTLT 1171

Qy 56 VNEDEKSIQAEFLFMNADNMGPATEKVTAVMYGYNRENAFRYETEDGVFTDVAIS 115

Db 1172 INED-KVWRADF-----AD-----DKIEDIKRYVNDLEIYQSF-----GYT 1208

Qy 116 DDNCDDVYVFTDGNNEGEYELMTTDYDNIIPANC-LNKFNVAVGRTRDVF 165

Db 1209 NSNGEQLWVRNAQ--NEBAHAV---FRNLSSDNNNSKYGSLSVDQINAF 1253

Search completed: August 1, 2003, 12:16:34

Job time: 14.7029 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using sw model

Run on: August 1, 2003, 12:15:58 ; Search time 16.9371 Seconds  
(without alignments)  
1199.017 Million cell updates/sec

Title: US-10-087-195-7  
Perfect score: 936  
Sequence: 1 NOPMDAANGAHQDAMKS.....FNEYAVGRETDRVTSACLE 171

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCF\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCFUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	936	100.0	171	US-10-085-572-7	Sequence 7, Appl 1
2	936	100.0	171	US-10-087-195-7	Sequence 7, Appl 1
3	630	67.3	172	US-10-085-572-6	Sequence 6, Appl 1
4	630	67.3	172	US-10-087-195-6	Sequence 6, Appl 1
5	331.5	35.4	182	US-10-085-572-8	Sequence 8, Appl 1
6	331.5	35.4	182	US-10-087-195-8	Sequence 8, Appl 1
7	91.5	9.8	221	US-09-728-914-4	Sequence 4, Appl 1
8	80	8.5	249	US-10-278-173-2	Sequence 22, Appl 1
9	80	8.5	329	US-09-738-626-5144	Sequence 5144, Ap
10	79	8.4	327	US-09-815-242-5484	Sequence 5484, Ap
11	79	8.4	327	US-09-815-242-12586	Sequence 12586, A
12	78.5	8.4	241	US-09-805-354-17	Sequence 17, Appl 1
13	78.5	8.4	241	US-10-144-259-17	Sequence 17, Appl 1
14	76.5	8.2	841	US-09-815-242-5779	Sequence 5779, Ap
15	76.5	8.2	841	US-09-815-242-12751	Sequence 12751, A

16	75	8.0	624	US-10-128-714-3284	Sequence 3284, Ap
17	75	8.0	843	US-10-128-714-8284	Sequence 8284, Ap
18	75	8.0	1806	US-10-156-761-13509	Sequence 13509, A
19	73.5	7.9	695	US-09-784-554B-6	Sequence 6, Appl 1
20	72.5	7.7	602	US-10-195-158-5	Sequence 5, Appl 1
21	72.5	7.7	661	US-09-988-626-240	Sequence 240, App
22	72.5	7.7	661	US-09-988-687-240	Sequence 240, App
23	72.5	7.7	661	US-09-988-686-240	Sequence 240, App
24	72.5	7.7	932	US-09-815-242-5578	Sequence 5578, Ap
25	72.5	7.7	932	US-09-815-242-12438	Sequence 12438, A
26	72	7.7	603	US-10-161-403-106	Sequence 106, App
27	72	7.7	604	US-09-893-525-37	Sequence 37, Appl 1
28	72	7.7	659	US-09-893-525-40	Sequence 40, Appl 1
29	72	7.7	743	US-10-087-464-53	Sequence 53, Appl 1
30	72	7.7	850	US-09-893-525-42	Sequence 42, Appl 1
31	72	7.7	1010	US-09-118-276-12	Sequence 12, Appl 1
32	71.5	7.6	444	US-10-112-488-40	Sequence 40, Appl 1
33	71.5	7.6	477	US-10-112-488-42	Sequence 42, Appl 1
34	71.5	7.6	878	US-09-912-020-364	Sequence 364, App
35	71.5	7.6	1198	US-09-975-719-405	Sequence 405, App
36	71	7.6	535	US-09-988-200-4	Sequence 4, Appl 1
37	71	7.6	1350	US-09-784-554B-4	Sequence 4, Appl 1
38	71	7.6	1956	US-10-209-776-2	Sequence 2, Appl 1
39	71	7.6	2764	US-09-808-602-80	Sequence 80, Appl 1
40	71	7.6	2764	US-09-800-198-68	Sequence 68, Appl 1
41	71	7.6	2765	US-09-808-602-84	Sequence 84, Appl 1
42	71	7.6	2765	US-09-800-198-72	Sequence 72, Appl 1
43	70.5	7.5	465	US-10-081-872-176	Sequence 176, App
44	70.5	7.5	2802	US-09-808-602-81	Sequence 81, Appl 1
45	70.5	7.5	2802	US-09-800-198-69	Sequence 69, Appl 1

## ALIGNMENTS

RESULT 1					
US-10-085-572-7					
Sequence 7, Application US/10085572					
Publication No. US20020151499A1					
GENERAL INFORMATION:					
APPLICANT: Nuttall, Patricia, Ann					
APPLICANT: Paesen, Guido, Christiaan					
TITLE OF INVENTION: Treatment of Conjunctionitis					
FILE REFERENCE: 2488-1-003					
CURRENT APPLICATION NUMBER: US/10/085,572					
CURRENT FILING DATE: 2002-02-27					
PRIOR APPLICATION NUMBER: PCT/GB00/03282					
PRIOR FILING DATE: 2000-08-24					
PRIOR APPLICATION NUMBER: 9920674.0					
PRIOR FILING DATE: 1999-09-01					
NUMBER OF SEQ ID NOS: 8					
SOFTWARE: FastSeq for Windows Version 4.0					
SEQ ID NO 7					
LENGTH: 171					
TYPE: PRT					
ORGANISM: FS-HBP 2					
US-10-085-572-7					
Query Match					
Best Local Similarity 100.0% ; Score 936; DB 14; Length 171;					
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	NOPMDAANGAHQDAMKS	KADVENYVYVYKATKKNDPVGNGFTCVGVANDVND	60	
DB	1	NOPMDAANGAHQDAMKS	LKADVENYVYVYKATKKNDPVGNGFTCVGVANDVND	60	
QY	61	KSIQAEFLMNNADTNM	FATEKYAVVYGVYGNREBNARYETEDQVETDVIYASDDNC	120	
DB	61	KSIQAEFLMNNADTNM	FATEKYAVVYGVYGNREBNARYETEDQVETDVIYASDDNC	120	
QY	121	VIVYVGTGNEGEYELMT	TDYDNIIPANLKNFNEAVGRETDRVTSACLE	171	
DB	121	VIVYVGTGNEGEYELMT	TDYDNIIPANLKNFNEAVGRETDRVTSACLE	171	

RESULT 2  
US-10-087-195-7  
Sequence 7, Application US/10087195  
Publication No. US20020193306A1  
GENERAL INFORMATION:  
APPLICANT: Nuttall, Patricia, Anne  
APPLICANT: Paesen, Guido, Christiaan  
TITLE OF INVENTION: Treatment of Allergic Rhinitis  
FILE REFERENCE: 2488-1-004  
CURRENT APPLICATION NUMBER: US/10/087,195  
CURRENT FILING DATE: 2002-03-01  
PRIOR APPLICATION NUMBER: PCT/GB00/03287  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: 9920673.2  
PRIOR FILING DATE: 1999-09-01  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 7  
LENGTH: 171  
TYPE: PRT  
ORGANISM: FS-HBP 2  
US-10-087-195-7

Query Match 100.0%; Score 936; DB 14; Length 171;  
Best Local Similarity 100.0%; Pred. No. 1.1e-92;  
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NOPDMADEAANGAHODAMKSLKADVENYVMKATYKNDPVWGNDFTCVGMANDVNEDE 60  
DB 1 NOPDMADEAANGAHODAMKSLKADVENYVMKATYKNDPVWGNDFTCVGMANDVNEDE 60  
QY 61 KSIQAEFLFNNMADTNNOFATEKYAVKMGYNRENAFRYETEDGOVFTVIASDNCND 120  
DB 61 KSIQAEFLFNNMADTNNOFATEKYAVKMGYNRENAFRYETEDGOVFTVIASDNCND 120  
QY 121 VIYVPGTGDGNEGEYELMTTIDYDNIIPANCLKFNENYAVGRETRDVFYSACLE 171  
DB 121 VIYVPGTGDGNEGEYELMTTIDYDNIIPANCLKFNENYAVGRETRDVFYSACLE 171

RESULT 3  
US-10-085-572-6  
Sequence 6, Application US/10085572  
Publication No. US20020151499A1  
GENERAL INFORMATION:  
APPLICANT: Nuttall, Patricia, Ann  
APPLICANT: Paesen, Guido, Christiaan  
TITLE OF INVENTION: Treatment of Conjunctivitis  
FILE REFERENCE: 2488-1-003  
CURRENT APPLICATION NUMBER: US/10/085,572  
CURRENT FILING DATE: 2002-02-27  
PRIOR APPLICATION NUMBER: PCT/GB00/03282  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: 9920674.0  
PRIOR FILING DATE: 1999-09-01  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 172  
TYPE: PRT  
ORGANISM: FS-HBP1  
US-10-085-572-6

Query Match 67.3%; Score 630; DB 14; Length 172;  
Best Local Similarity 66.5%; Pred. No. 7.4e-60;  
Matches 113; Conservative 20; Mismatches 37; Indels 0; Gaps 0;

QY 1 NOPDMADEAANGAHODAMKSLKADVENYVMKATYKNDPVWGNDFTCVGMANDVNEDE 60  
DB 1 KRPVMADEAANGAHODAMKSLKADVENYVMKATYKNDPVWGNDFTCVGMANDVNEDE 60

QY 61 KSIQAEFLFNNMADTNNOFATEKYAVKMGYNRENAFRYETEDGOVFTVIASDNCND 120  
DB 61 KNEVAMFEMNNMADTYQHTEKATPDKMGYNKENAITYQTEDEGVLTDLAFSDNCY 120  
QY 121 VIYVPGTGDGNEGEYELMTTIDYDNIIPANCLKFNENYAVGRETRDVFYSACLE 170  
DB 121 VIYVPGTGDGAGAGYELMATDYTDVPASCLKFNENYAGLVRDVTSDCL 170

RESULT 4  
US-10-087-195-6  
Sequence 6, Application US/10087195  
Publication No. US20020193306A1  
GENERAL INFORMATION:  
APPLICANT: Nuttall, Patricia, Anne  
APPLICANT: Paesen, Guido, Christiaan  
TITLE OF INVENTION: Treatment of Allergic Rhinitis  
FILE REFERENCE: 2488-1-004  
CURRENT APPLICATION NUMBER: US/10/087,195  
CURRENT FILING DATE: 2002-03-01  
PRIOR APPLICATION NUMBER: PCT/GB00/03287  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: 9920673.2  
PRIOR FILING DATE: 1999-09-01  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 172  
TYPE: PRT  
ORGANISM: FS-HBP1  
US-10-087-195-6

Query Match 67.3%; Score 630; DB 14; Length 172;  
Best Local Similarity 66.5%; Pred. No. 7.4e-60;  
Matches 113; Conservative 20; Mismatches 37; Indels 0; Gaps 0;

QY 1 NOPDMADEAANGAHODAMKSLKADVENYVMKATYKNDPVWGNDFTCVGMANDVNEDE 60  
DB 1 KRPVMADEAANGAHODAMKSLKADVENYVMKATYKNDPVWGNDFTCVGMANDVNEDE 60  
QY 61 KSIQAEFLFNNMADTNNOFATEKYAVKMGYNRENAFRYETEDGOVFTVIASDNCND 120  
DB 61 KNEVAMFEMNNMADTYQHTEKATPDKMGYNKENAITYQTEDEGVLTDLAFSDNCY 120  
QY 121 VIYVPGTGDGNEGEYELMTTIDYDNIIPANCLKFNENYAVGRETRDVFYSACLE 170  
DB 121 VIYVPGTGDGAGAGYELMATDYTDVPASCLKFNENYAGLVRDVTSDCL 170

RESULT 5  
US-10-085-572-8  
Sequence 8, Application US/10085572  
Publication No. US20020151499A1  
GENERAL INFORMATION:  
APPLICANT: Nuttall, Patricia, Ann  
APPLICANT: Paesen, Guido, Christiaan  
TITLE OF INVENTION: Treatment of Conjunctivitis  
FILE REFERENCE: 2488-1-003  
CURRENT APPLICATION NUMBER: US/10/085,572  
CURRENT FILING DATE: 2002-02-27  
PRIOR APPLICATION NUMBER: PCT/GB00/03282  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: 9920674.0  
PRIOR FILING DATE: 1999-09-01  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8  
LENGTH: 182  
TYPE: PRT  
ORGANISM: MS-HBP 1  
US-10-085-572-8

Query Match 35.4%; Score 331.5; DB 14; Length 182;

Best Local Similarity 42.0%; Pred. No. 8.6e-28;  
Matches 74; Conservative 29; Mismatches 62; Indels 11; Gaps 7;

Oy 3 PDMADEANGAHODAMSKLADVENYVYKATYKNDPVGNDFTCGVANDVNEDEKS 62  
Db 2 PTMANEAKLSYODAMSLDODONKRYLLAOATOTGTGNGEFTCVSAEKG--KKK 59

Oy 63 IOAEFLPMNADTNMOPATEKRYAVKMYGNRENAFYETED--GOVFTDVIASD-DNC 119  
Db 60 LNAITILKNNHLDTLKSHEITVYWKAYDYTTENGIRYETGTGTQTFEDVFEVSDYKNC 119

Oy 120 DVIVPFGTGNESG-YELMTTD--YDNIIPANCLKFNENYAGRE--TRDVF--SAC 169  
Db 120 DVIFVPERKSGDESDYELMWSEDKIDIPDCKFTMAVFAOQDEKTVRNVYTTDSSC 175

RESULT 6  
US-10-087-195-8  
Sequence 8, Application US/10087195  
Publication No. US20020193306A1

GENERAL INFORMATION:  
APPLICANT: Nuttall, Patricia, Anne  
TITLE OF INVENTION: Treatment of Allergic Rhinitis  
FILE REFERENCE: 2488-1-004  
CURRENT APPLICATION NUMBER: US/10/087,195  
PRIOR APPLICATION NUMBER: PCT/GB00/03287  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: 9920673.2  
PRIOR FILING DATE: 1999-09-01  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8

LENGTH: 182  
TYPE: PRT  
ORGANISM: MS-HBP 1  
US-10-087-195-8

Query Match  
Best Local Similarity 42.0%; Pred. No. 8.6e-28;  
Matches 74; Conservative 29; Mismatches 62; Indels 11; Gaps 7;

Oy 3 PDMADEANGAHODAMSKLADVENYVYKATYKNDPVGNDFTCGVANDVNEDEKS 62  
Db 2 PTMANEAKLSYODAMSLDODONKRYLLAOATOTGTGNGEFTCVSAEKG--KKK 59

Oy 63 IOAEFLPMNADTNMOPATEKRYAVKMYGNRENAFYETED--GOVFTDVIASD-DNC 119  
Db 60 LNAITILKNNHLDTLKSHEITVYWKAYDYTTENGIRYETGTGTQTFEDVFEVSDYKNC 119

Oy 120 DVIVPFGTGNESG-YELMTTD--YDNIIPANCLKFNENYAGRE--TRDVF--SAC 169  
Db 120 DVIFVPERKSGDESDYELMWSEDKIDIPDCKFTMAVFAOQDEKTVRNVYTTDSSC 175

RESULT 7  
US-09-728-914-4  
Sequence 4, Application US/09728914  
Patent No. US20010046499A1  
GENERAL INFORMATION:  
APPLICANT: KANTOR, FRED S.  
APPLICANT: FIKRIG, EROL  
APPLICANT: DAS, SUBRATA  
TITLE OF INVENTION: TICK ANTIGENS AND COMPOSITIONS AND METHODS COMPRISING  
FILE REFERENCE: YU-107  
CURRENT APPLICATION NUMBER: US/09/728,914  
CURRENT FILING DATE: 2000-12-01  
PRIOR APPLICATION NUMBER: 60/169,048  
PRIOR FILING DATE: 1999-12-03  
PRIOR APPLICATION NUMBER: 60/240,716  
PRIOR FILING DATE: 2000-10-16

NUMBER OF SEQ ID NOS: 39  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 4  
LENGTH: 221  
TYPE: PRT  
ORGANISM: Ixodes scapularis  
US-09-728-914-4

Query Match  
Best Local Similarity 24.2%; Pred. No. 0.063;  
Matches 32; Conservative 24; Mismatches 69; Indels 7; Gaps 5;

Oy 45 DETCGVANDVNEDEKSIOAEFLPMNADTNMOPAT-EKYAVKMYG-YNRENAFYET 102  
Db 88 DFGSNPVOECENFRMERKTRPNISFOYRYKSKNSMETIDETLLIKDIDGEGFPVNMFOR 147

Oy 103 EDGOVFTD-VIAYS-DNCDVYVPGTGNESG-YELMTTD--DNIIPANCLKFNENYAV 157  
Db 148 TPGIATDNLVLYSNVNCVLRIPFTNOGBRCHDLWMAN/LTISOETPDCLNKFFECYN 207

Oy 158 GRETRDVFYSAC 169  
Db 208 TQIYRYVYPSG 219

RESULT 8  
US-10-278-173-22  
Sequence 22, Application US/10278173  
Publication No. US20030061637A1  
GENERAL INFORMATION:  
APPLICANT: Jiang, Cai-Zhong  
APPLICANT: Brown, Pierre  
APPLICANT: Riechmann, Jose-Luis  
APPLICANT: Pineda, Omaira  
APPLICANT: Zhang, James  
APPLICANT: Yu, Guo-Liang  
APPLICANT: Pilgrim, Marsha  
APPLICANT: Keddie, James  
APPLICANT: Heard, Jacqueline  
APPLICANT: Reuber, Lynne  
APPLICANT: Ratcliffe, Oliver  
APPLICANT: Adam, Luc  
APPLICANT: Samaha, Raymond  
TITLE OF INVENTION: POLYNUCLEOTIDES FOR ROOT TRAIT ALTERATION  
FILE REFERENCE: MBI-009  
CURRENT APPLICATION NUMBER: US/10/278,173  
PRIOR APPLICATION NUMBER: 2002-10-21  
PRIOR FILING DATE: 2000-03-22  
PRIOR APPLICATION NUMBER: 60/125,814  
PRIOR FILING DATE: 1999-03-23  
NUMBER OF SEQ ID NOS: 177  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 22  
LENGTH: 249  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
OTHER INFORMATION: G1307  
US-10-278-173-22

Query Match  
Best Local Similarity 22.6%; Pred. No. 1.3;  
Matches 38; Conservative 27; Mismatches 65; Indels 38; Gaps 7;

Oy 18 WKSJKA-----DVENVYVYKATYKNDPVGNDFTCGVANDVNEDEKSIOAEFL 68  
Db 89 WSAIAAKLPGRDNEIKNVHTHLKRLSKNLNNGGPTKDVNGINETTNEDKGSVIYD- 146

Oy 69 FMNADTNMOPATEKRYAVKMYGNRENAFYETEDGOVFTDVIASD--CVIYVPG 126  
Db 147 -----TASLOQFSNSTITFTDISDNKDIMSVE-----DISALIDDSFWSVSIYDN 193

QY 127. TDGNESEYELMTDYDNIPANC-----LNKNEVAVGRETVDFTS 167  
Db 194. SNRNEKKEIDMEGLIDRNSKRCYSNSKLYNDMEFVF-----DVFTS 236

## RESULT 9

US-09-738-626-5144  
; Sequence 5144, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIRO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAOKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 5144  
; LENGTH: 929  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-5144

Query Match 8.5%; Score 80; DB 10; Length 929;  
Best Local Similarity 25.6%; Pred. No. 7.4;  
Matches 31; Conservative 20; Mismatches 56; Indels 14; Gaps 3;

QY 21 LKADVENVYVAVATKNDPWCNDTCGVAMANDV--NDEKSIOAEFLPMNADTNMQ 78  
Db 73 IKALSNOKYHDIAKAGSD-----AVGLITGDVSNINDADIVVWTEVLRNMIYAGS 124  
QY 79 FATEKYAVVWYGVNRENAFRYETEDGOVFTDVIASDNCQVIVYPGTDCNDEGEYELWT 138  
Db 125 FALERLSHYV-----DEIHFLADASKGAVWEVYILNLDVSNITIGLSATVSNSEFEGL 180  
QY 139 T 139  
Db 181 T 181

## RESULT 10

US-09-815-242-5484  
; Sequence 5484, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21

;; PRIOR APPLICATION NUMBER: 60/191,078  
;; PRIOR FILING DATE: 2000-03-21  
;; PRIOR APPLICATION NUMBER: 60/206,848  
;; PRIOR FILING DATE: 2000-05-23  
;; PRIOR APPLICATION NUMBER: 60/207,727  
;; PRIOR FILING DATE: 2000-05-26  
;; PRIOR APPLICATION NUMBER: 60/242,578  
;; PRIOR FILING DATE: 2000-10-23  
;; PRIOR APPLICATION NUMBER: 60/253,625  
;; PRIOR FILING DATE: 2000-11-27  
;; PRIOR APPLICATION NUMBER: 60/257,931  
;; PRIOR FILING DATE: 2000-12-22  
;; PRIOR APPLICATION NUMBER: 60/269,308  
;; PRIOR FILING DATE: 2001-02-16  
;; NUMBER OF SEQ ID NOS: 1410  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 5484  
;; LENGTH: 327  
;; TYPE: PRT  
;; ORGANISM: Staphylococcus aureus  
US-09-815-242-5484

Query Match 8.4%; Score 79; DB 9; Length 327;  
Best Local Similarity 23.1%; Pred. No. 2.4;  
Matches 39; Conservative 25; Mismatches 63; Indels 42; Gaps 8;

QY 14 HODAMSKADVENVYVAV--KATYKNDPWCNDTC--CVGVAMANDNEKSIOAEFL 68  
Db 173 HKRAYRKEVEVEEYTYVPLGKADYKRE---GEDLTVPYGLAMVYICLQAAIILAADI 229  
QY 69 FMNADTNMQFATEKYVAV-----KMGYNRENAFRYETEDGOVFTDVIASDNC-- 119  
Db 230 NVEVYDLRYVYPLDKETIIDRAKHGKVLVTFEDNL-----GSIINSEVSAILIARHCLF 283  
QY 120 -----DVIYVPGTDCNDEGEYELMTTYDNIIPANCUKKFNEXA 156  
Db 284 ELDAPIRLAAPVPSMPSPVLEN--ETIMN-----PEKILNKRRELA 325

## RESULT 11

US-09-815-242-12586  
; Sequence 12586, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 1410  
; SOFTWARE: FastSeq for Windows Version 4.0

```

; SEQ ID NO 12586
;
; LENGTH: 327
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
;
US-09-815-242-12586

```

Query Match	8.48;	Score 79;	DB 9;	Length 327;
Best Local Similarity	23.18;	Pred. NO. 2.4;		
Matches 39; Conservative	25;	Mismatches 63;	Indels 42;	Gaps 8

```

0Y      14 HQAMKSLADYENYYNY---KATYKNDPVGNFT---CYGVAMNDYNEKSIQAEFL 68
      | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db      173 HKAYAFLEEVPEEYTYPLGKADYKE---GEDLYFCGLMANNCLAADILANDGI 229
      | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
0Y      69 FMANNADTNQFATEKYTAV-----KMGYRENAFRYETEDGOVFTDLYAYSDDNC-- 119

```

Db 230 NWEVVDLRTVYPLDKETIIDRAKHGKVLVLTEDNL-----EGSINSEVSALIAEHCLF 283

Qy 120 -----DVIYVPGTDCNGEGVELMTTDDYDNIPANCLANKFNEYA 156  
|| : : : : ||| |

Db 284 ELDAPIMRLAAPVPSMPSVLEN--EIMNN-----DEKILNKRELA 325

RESULT 12  
US-09-805-354-17  
; Sequence 17, Application US/09805354  
; Publication No. US20030078375A1

```

: APPLICANT: Xiong, Jian-Ping
: TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND USES THEREOF
: FILE REFERENCE: 00786-536001
: CURRENT APPLICATION NUMBER: US/09/805,354

```

```

; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0

```

ORGANISM: Homo sapiens  
US-09-805-354-17

Query Match	8.48;	Score 78.5;	DB 11;	Length 241;
Best Local Similarity	23.18;	Pred. No. 1.8;		
Matches 28; Conservative	20;	Mismatches 50;	Indels 23;	Gaps 5

```

OY 42 MGN-----DFTCYGVANOVNEDKSIQAELEFFMN-----NADTNMFAF-----EKV 84
Db 67 WPNSDPEPFSKKNVISTLEDYDEFKRNKLGQERISGNLDAPEGGFDAIIQTAVCTRDIGMRP 1262
OY 85 TAYKMYGVNENMARETEDQVFPYDVAAYSPNDODVYYVGTGDNEGEYELMTYDNDI 1444
Db 127 DSHLLVSTESNPHETLEADGANVLGINSRNDERCHL-----DFTGYTYQY--TQDPSPV 1800

```

Qy	145 P 145
	1
Db	181 P 181

RESULT 13  
US-10-144-259-17  
; Sequence 17, Application US/10144259  
; Publication No. US20030109691A1.

```

: APPLICANT: Arnaud, M. Amin
:
: APPLICANT: Li, Rui
:
: APPLICANT: Xiong, Jian-Ping
:
: TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF

```

```

: FILE REFERENCE: 00786-548001
: CURRENT APPLICATION NUMBER: US/10/144,259
: CURRENT FILING DATE: 2002-09-04
: PRIOR APPLICATION NUMBER: US 09/758,493
: PRIOR FILING DATE: 2001-01-11
: PRIOR APPLICATION NUMBER: US 60/221,950
: PRIOR FILING DATE: 2000-07-31
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: FastSeq for Windows Version 4.0.
: SEQ ID NO: 17
: LENGTH: 241
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-144-259-17

```

ORGANISM: Homo sapiens  
US-10-144-259-17

Query Match	8.4%;	Score 78.5;	DB 15;	Length 241;
Best Local Similarity	23.1%;	Pred. No. 1.8;		
Matches	28;	Conservative	20;	Mismatches 50;
				Indels 23;
				Gaps 5

QY 42 WGN-----DFTCVGVAMANDVNEDEKSIQAEFLFMN-----NADTNMQEAT-----EKV 84

QY 85 TAVNMGYNKRNENAFRYETEDGQVFETDIAYISDDNCDIYYIPGTGNEGELMTTDDYDNI 144  
:  
Db 127 DSTHLLVFTSESAFHYEADGANVLGIMSRDERCHL---DTGTGYTYQR--TDQVPSPV 180

QY	145 P 145
	—
Db	181 P 181

RESULT 14  
US-09-815-242-5779  
; Sequence 5779, Application US/09815242  
; Patent No. US20020061569A1

```

: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlssen, Karl L.
: APPLICANT: Zykkind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes in
: TITLE OF INVENTION: Prokaryotes

```

CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21

; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,722  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23

; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22

```

; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5779
; LENGTH: 841
;

```

```

: ORGANISM: Staphylococcus aureus
US-09-815-242-5779

Query Match      8.2% Score 76.5; DB 9; Length 841

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: August 1, 2003, 12:15:04 ; Search time 13.0286 seconds  
(without alignments)  
555.330 Million cell updates/sec

Title: US-10-087-195-7

Perfect score: 936  
Sequence: 1 NQPDMAEANGAHQDAWKS.....FNEYAVGRETRDFVSACLE 171

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/1aa/PTUS.COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/1aa/Backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	85.5	9.1	337	US-09-134-001C-3207	Sequence 3207, Ap
2	82.5	8.8	405	US-08-282-197C-61	Sequence 61, Appl
3	78.5	8.4	252	PCR-US96-01314-56	Sequence 56, Appl
4	76.5	8.2	489	PCR-US93-11110-1	Sequence 1, Appl
5	75.5	8.1	259	US-08-871-572B-6	Sequence 6, Appl
6	75.5	8.1	489	US-08-190-204-2	Sequence 2, Appl
7	75.5	8.1	489	5221789-1	Patent No. 5221789
8	74.5	8.0	1964	US-08-790-912-3	Sequence 3, Appl
9	74.5	8.0	2052	US-08-790-912-2	Sequence 2, Appl
10	74	7.9	420	US-09-107-533A-5946	Sequence 5946, Ap
11	73.5	7.9	113	US-09-107-533A-6865	Sequence 6865, Ap
12	73.5	7.9	489	US-09-107-533A-4711	Sequence 4711, Ap
13	73	7.8	1335	US-09-134-001C-4716	Sequence 3716, Ap
14	72.5	7.7	600	5268463-2	Patent No. 5268463
15	72.5	7.7	602	US-08-882-704A-5	Sequence 5, Appl
16	72.5	7.7	602	US-09-151-957-5	Sequence 5, Appl
17	72.5	7.7	602	5432081-2	Patent No. 5432081
18	72.5	7.7	661	US-09-564-805-240	Sequence 240, App
19	72.5	7.7	1013	US-09-062-126-3	Sequence 126-3
20	72.5	7.7	1022	US-08-772-270A-2	Sequence 2, Appl
21	72.5	7.7	1242	US-09-488-270A-2	Sequence 2, Appl
22	72.5	7.7	1956	US-08-843-417-10	Sequence 10, Appl
23	72.5	7.7	1956	US-09-527-013-10	Sequence 10, Appl
24	72	7.7	329	US-09-071-035-368	Sequence 368, App
25	72	7.7	357	US-09-071-035-368	Sequence 366, App
26	72	7.7	603	US-09-149-727-6	Sequence 6, Appl
27	72	7.7	740	US-08-257-073-5	Sequence 5, Appl

28	72	7.7	832	3	US-08-630-820-7	Sequence 7, Appl
29	72	7.7	1227	1	US-08-448-170-8	Sequence 8, Appl
30	72	7.7	1227	3	US-08-961-803-9	Sequence 9, Appl
31	71.5	7.6	611	1	US-08-386-727-4	Sequence 4, Appl
32	71.5	7.6	611	2	US-08-600-452A-4	Sequence 4, Appl
33	71.5	7.6	1198	4	US-09-199-637A-405	Sequence 405, App
34	71	7.6	223	3	US-09-254-733-7	Sequence 7, Appl
35	71	7.6	347	4	US-09-328-353-6764	Sequence 6764, Ap
36	71	7.6	535	4	US-09-269-731-4	Sequence 4, Appl
37	71	7.6	1174	1	US-08-040-751-3	Sequence 3, Appl
38	71	7.6	1174	1	US-08-291-368-2	Sequence 2, Appl
39	71	7.6	1174	2	US-08-962-190-2	Sequence 2, Appl
40	71	7.6	1174	5	PCR-US95-10310-2	Sequence 2, Appl
41	71	7.6	1174	6	5164180-4	Patent No. 5164180
42	71	7.6	1229	1	US-08-100-709-4	Sequence 4, Appl
43	71	7.6	1229	1	US-08-176-865-4	Sequence 4, Appl
44	71	7.6	1229	1	US-08-474-038-4	Sequence 4, Appl
45	71	7.6	1229	2	US-08-779-046-4	Sequence 4, Appl

## ALIGNMENTS

```
RESULT 1
US-09-134-001C-3207
; Sequence 3207, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3207
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3207

Query Match          9.1%; Score 85.5; DB 4; Length 337;
Best Local Similarity 27.5%; Pred. No. 0.29;
Matches 33; Conservative 16; Mismatches 46; Indels 25; Gaps 7;

QY      14 HODAKSLKADYENYVWY---KATYKNDPVWGNDFT---CVGVMANDVDEKSTOAEFL 68
      183 HKKATRFKEEYEPETTYVPLCKADYKRP---GEDITVFCYGLMVA-----YCLQADI 233
      69 FNNV-----ADTNQFATEKYAVAMGYNRENAFRETED---GOVTDVIYASDDNC 119
      234 LANDGDIVEVDLRFYVPLDKATIER--SQRTGKVLVTEQNLBSINSISVSAITAENC 291

RESULT 2
US-08-282-197C-61
; Sequence 61, Application US/08282197C
; Patent No. 5871730
; GENERAL INFORMATION:
; APPLICANT: Brzezinski, Ryszard
; APPLICANT: Dery, Claude V
; APPLICANT: Beaulieu, Carole
; TITLE OF INVENTION: Thermostable Xylanase DNA, Protein and
; TITLE OF INVENTION: Methods of Use
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Ave., NW
; CITY: Washington
```

Query Match	8.8%	Score	82.5	DB	2	Length	405
Best Local Similarity	24.3%	Pred	No	0.82			
Matches	43	Conservative	26	Mismatches	37	Indels	71
						Gaps	13

RESULT 3  
PCT-US96-01314-56

APPLICANT: M. Amin Annaout  
TITLE OF INVENTION: METHODS FOR IDENTIFYING  
TITLE OF INVENTION: ANTAONISTS  
NUMBER OF INVENTIONS: 78  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U S.A  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
COMPUTER: IBM PS/2 Model 50Z or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: Wordperfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/01314  
FILING DATE: 30-JAN-96  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/380,167  
FILING DATE: 30-JAN-95  
ATTORNEY/AGENT INFORMATION:  
NAME: John W. Freeman  
REGISTRATION NUMBER: 29,066  
REFERENCE/DOCKET NUMBER: 00786/267001

Query Match	8.48	Score	78.5	DB	5	Length	252
Best Local Similarity	23.18	Pred. No.	1.1				
Matches	28	Conservative	20	Mismatches	50	Indels	23
						Gaps	5

QY	145 P 145
Db	196 P 196

RESULT 4  
PCT-US93-11110-1  
; Sequence 1, Application PC/TUS9311110

```

1  APPLICANT: Schering Corp.
2  TITLE OF INVENTION: Antagonists of Human Gamma Interferon
3  NUMBER OF SEQUENCES: 8
4  CORRESPONDENCE ADDRESS:
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
```

;  
; CITY: Madison  
;  
; STATE: New Jersey  
;  
; COUNTRY: USA  
;  
;

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh 6.0.8  
SOFTWARE: Microsoft Word 5.1a  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/11110  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA: U.S. Patent Application No. 07/980,527 filed November 2  
ATTORNEY/AGENT INFORMATION:  
NAME: Lunn, Paul G.  
REGISTRATION NUMBER: 32,743  
REFERENCE/DOCKET NUMBER: JB0285Q  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-822-7255  
TELEFAX: 201-822-7039  
TELEX: 219165  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 489 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US93-11110-1

Query Match	8.2%	Score 76.5;	DB 5;	Length 489;
Best Local Similarity	19.3%	Pred. No. 5;		
Matches 37; Conservative	30;	Mismatches 70;	Indels 55;	Gaps 11

OY 4 DWADEANGAH-----ODAKSLKADY---ENVYVVKATYKNDPVWGND- 46  
DB 72 EMIDACINISHHCNISDHVGDPSNSLWVRKARVGOKESAY-----AKSEFEA 120  
OY 47 TC-----VGVMANDVNEDEKSIQAEF---LFMNNADTMQFATEKVTAVKMG-YNRENA 97  
DB 121 VCRGKIGPRLDIRKEKQIMIDIFHPSVFNVDGDEVDYDPEETCYIRVYNYVRNG 180  
OY 98 FRYETEDGQVFTDVIASDDNCVYI---VPGTDGNE-----EG-YELMTTDDNIPA 146  
DB 181 -----SEIOYKILTKQEDDCDEIQCOLAIPVSLNSQYCSAEGVLHVGVTTEKSKE 233  
OY 147 NCLKNFNEYAVG 158  
DB 234 VCITTFNSSING 245

RESULT 5  
US-08-871-572B-6  
Sequence 6, Application US/08871572B  
Patent No. 6287853  
GENERAL INFORMATION:  
APPLICANT: Pestka, Sidney  
APPLICANT: Kolenko, Serguei  
APPLICANT: Soh, Jaemog  
APPLICANT: Donnelly, Robert  
APPLICANT: Mariano, Thomas  
APPLICANT: Cook, Jeffrey  
APPLICANT: Emmanuel, Stuart  
APPLICANT: Schwartz, Barbara  
TITLE OF INVENTION: Accessory Factor for Interferon Gamma  
TITLE OF INVENTION: and its Receptor  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Richard R. Muccino  
STREET: 758 Springfield Avenue  
CITY: Summit  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07901  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/871,572B  
FILING DATE: 9-JUNE-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Muccino, Richard R.  
REGISTRATION NUMBER: 32,538  
REFERENCE/DOCKET NUMBER: UMD1-011  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 273-4988  
TELEFAX: (908) 273-4679  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 259 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
US-08-871-572B-6

Query Match 8.1%; Score 75.5; DB 3; Length 259;  
Best Local Similarity 19.3%; Pred. No. 2.5; Indels 55; Gaps 11;  
Matches 37; Conservative 29; Mismatches 71; Indels 55; Gaps 11;  
4 DWADEANGAH-----ODAKSLKADY---ENVYVVKATYKNDPVWGND- 46

DB 72 EMIDACINISHHCNISDHVGDPSNSLWVRKARVGOKESAY-----AKSEFEA 120  
OY 47 TC-----VGVMANDVNEDEKSIQAEF---LFMNNADTMQFATEKVTAVKMG-YNRENA 97  
DB 121 VCRGKIGPRLDIRKEKQIMIDIFHPSVFNVDGDEVDYDPEETCYIRVYNYVRNG 180  
OY 98 FRYETEDGQVFTDVIASDDNCVYI---VPGTDGNE-----EG-YELMTTDDNIPA 146  
DB 181 -----SEIOYKILTKQEDDCDEIQCOLAIPVSLNSQYCSAEGVLHVGVTTEKSKE 233  
OY 147 NCLKNFNEYAVG 158  
DB 234 VCITTFNSSING 245

RESULT 6  
US-08-190-204-2  
Sequence 2, Application US/08190204  
Patent No. 6558661  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Ward, Rebecca H.R.  
TITLE OF INVENTION: TREATMENT OF INFLAMMATORY BOWEL DISEASE WITH IFN-GAMMA  
TITLE OF INVENTION: INHIBITORS  
FILE REFERENCE: 11669, 90USMO  
CURRENT APPLICATION NUMBER: US/08/190,204  
CURRENT FILING DATE: 1994-02-22  
PRIOR APPLICATION NUMBER: 07/997,835  
PRIOR FILING DATE: 1992-12-29  
PRIOR APPLICATION NUMBER: PCT/US93/11966  
PRIOR FILING DATE: 1993-12-09  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 2  
LENGTH: 489  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-08-190-204-2

Query Match 8.1%; Score 75.5; DB 4; Length 489;  
Best Local Similarity 19.3%; Pred. No. 6.4;  
Matches 37; Conservative 29; Mismatches 71; Indels 55; Gaps 11;  
4 DWADEANGAH-----ODAKSLKADY---ENVYVVKATYKNDPVWGND- 46  
DB 72 EMIDACINISHHCNISDHVGDPSNSLWVRKARVGOKESAY-----AKSEFEA 120  
OY 47 TC-----VGVMANDVNEDEKSIQAEF---LFMNNADTMQFATEKVTAVKMG-YNRENA 97  
DB 121 VCRGKIGPRLDIRKEKQIMIDIFHPSVFNVDGDEVDYDPEETCYIRVYNYVRNG 180  
OY 98 FRYETEDGQVFTDVIASDDNCVYI---VPGTDGNE-----EG-YELMTTDDNIPA 146  
DB 181 -----SEIOYKILTKQEDDCDEIQCOLAIPVSLNSQYCSAEGVLHVGVTTEKSKE 233  
OY 147 NCLKNFNEYAVG 158  
DB 234 VCITTFNSSING 245

RESULT 7  
5221789-1  
Patent No. 5221789  
APPLICANT: NOVICK, DANIELA; RUBINSTEIN, MENACHEM  
TITLE OF INVENTION: INTERFERON-GAMMA RECEPTOR FRAGMENT AND  
ITS PRODUCTION  
NUMBER OF SEQUENCES: 1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/578,826  
FILING DATE: 07-SEP-1990  
SEQ ID NO: 1  
LENGTH: 489

5221789-1

Query Match	8.1%;	Score 75.5;	DB 6;	Length 489;
Best Local Similarity	19.3%;	Pred. No. 6.4;		
Matches 37;	Conservative 29;	Mismatches 71;	Indels 55;	Gaps 11.

```

QY      4 DAMEDEANGH-----QDAKSLKADY---ENYYMAKATYKNDPWWGDF- 46
Db      72 EWDIACINISHHYCNISDHVGDPSSMLWRYKAFYGOKE$AY-----AKSEFEA 120
        : : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY      47 TC-----VGWANDVNEDEKSIQAEF-----LENNADTMMOPATEKYTAVKMYG- YNSENA 97
Db      121 VCDGKIGIPKDKIDRIREEQIMIDLHFHSVFNPGEQGEVDYDDPETTCYIRKYNINYAMNG 180
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      98 FRFETEDGQYTFDVIAYSDNCQDVYI---VPGTDGNE-----EG-YEIMTDDYDNI 146
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      181 -----SEIQYKILQKEDDCDEIQOLAIIPVSSLNSQYCVSA$EGLHWGVTTESKE 233
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      147 NCLNKFNEYAYG 158
        : : : | : |
Db      234 VCIITFNSSIKG 245

```

RESULT 8  
ITS-08-79

Sequence 3, Application US/08790912  
Patent No. 5976542  
GENERAL INFORMATION:  
APPLICANT: Weiser, Jeffrey N.  
APPLICANT: Plaut, Andrew G.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT  
OF STREPTOCOCCUS PNEUMONIAE INFECTION  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PAATCH SCHWARZE JACOBS & NADEL, P.C.  
STREET: 1601 Market street, 36th floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103-2398  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/790,912  
FILING DATE: 29-JAN-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/026,939  
FILING DATE: 23-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Leaty, Kathryn  
REGISTRATION NUMBER: 36,317  
REFERENCE/DOCKET NUMBER: 7600-401  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 567-2020  
TELEFAX: (215) 567-2991  
TELEX: 831-494  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1964 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
IS-08-790-912-3

Query Match	8.0%;	Score 74.5;	DB 2;	Length 1964;
Best Local Similarity	25.2%;	Pred. No. 63;		
Matches 27;	Conservative 19;	Mismatches 32;	Indels 29;	Gaps 5;

```

OY 61 KSIAEFLFEMNADTNMQFATEKTYAVMYGYNNENAFRYETE-----DGOVFTEVI 112
Db 1124 KNIKKLLPEYN-----KSTIYKYGKLVKRENSLLYQKELLSAVMMKDDQYITDIY 1172
OY 113 AYSDDNDVLYVPGTDGNEEGVEL-WPTDYNINIRANCLNKRNKEPANG 158
Db 1173 S-NQOTAKKLLHYNDHSSEKFDLKTQTFDNP-----EYNLG 1210

```

## RESULT 9

```

1 Sequence 2, Application US/08790912
2 Patent No. 5976542
3
4 GENERAL INFORMATION:
5 APPLICANT: Weiser, Jeffrey N.
6 APPLICANT: Platt, Andrew G.
7 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT
8 OF STREPTOCOCCUS PNEUMONIAE INFECTION
9 NUMBER OF SEQUENCES: 10
10
11 CORRESPONDENCE ADDRESS:
12 ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
13 STREET: 1601 Market Street, 36th Floor
14 CITY: Philadelphia
15 STATE: Pennsylvania
16 COUNTRY: USA
17 ZIP: 19103-2398
18
19 COMPUTER READABLE FORM:
20 MEDIUM TYPE: Floppy disk
21 COMPUTER: IBM PC compatible
22 OPERATING SYSTEM: PC-DOS/MS-DOS
23 SOFTWARE: PatentIn Release #1.0, Version #1.30
24 CURRENT APPLICATION DATA:
25 APPLICATION NUMBER: US/08/790,912
26 FILING DATE: 29-JAN-1997
27 CLASSIFICATION: 424
28
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: US 60/026,939
31 FILING DATE: 23-SEP-1996
32 ATTORNEY/AGENT INFORMATION:
33 NAME: Leary, Kathryn
34 REGISTRATION NUMBER: 36,317
35 REFERENCE/DOCKET NUMBER: 7600-401
36 TELECOMMUNICATION INFORMATION:
37 TELEPHONE: (215) 567-2020
38 TELEFAX: (215) 567-2991
39
40 TELEEX: 831-494
41
42 INFORMATION FOR SEQ ID NO: 2:
43 SEQUENCE CHARACTERISTICS:
44 LENGTH: 2052 amino acids
45 TYPE: amino acid
46 TOPOLOGY: linear
47
48 MOLECULE TYPE: protein
49
50 US-08-790-912-2

```

Query Match	8.08;	Score 74.5;	DB 2;	Length 2052;
Best Local Similarity	25.28;	Pred. No. 67;		
Matches 27; Conservative	19;	Mismatches 32;	Indels 29;	Gaps 5;

```

OY      61 KSIQAEELFLPMNADTNMQPATEKXTAVKMGYNNENFRYTE-----DGOVETDVI 112
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1197 KNIEKLFLPFN-----KTIYKIGULVKNENSLIYKELLSANMMKDDQYITDVI 1245
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      113 AYSDNDNCVYIYVPGTDCNNEGIEL-WITTDYDNIPANCLINKENAYAVG 158
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      1246 S-NKQATANKLLHLHYNDHSSEKFDIKYOTDFANLP-----EKNLG 1283
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

## RESULT 10

US-03-107-032A-0940  
Sequence 5946, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Arinello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 5946:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 420 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (B) LOCATION 1...420  
SEQUENCE DESCRIPTION: SEQ ID NO: 5946:  
US-09-107-532A-5946  
Query Match 7.9%; Score 74; DB 4; Length 420;  
Best Local Similarity 23.3%; Pred. No. 7.5;  
Matches 41; Conservative 15; Mismatches 58; Indels 62; Gaps 9;  
QY 4 DWADDAANGAHQDAWKSIAKADVENYVYVAKATYKNDPYWGNDFTCVGYMANDV-----N 57  
DB 261 DWDD-----MLDVATNYPAPFETIKQDQDQGE-----GQMIDDIYQILGK 301  
QY 58 EDEKSIQAEF---LFMNNADTNMQPATEKTAIVAMGYKRNEMARFETEDGQVTFDVIAY 114  
DB 302 NDLSIONELKROKINIKSDTLPLKNE-----NASMFEFDSDEG----- 342  
QY 115 SDDNCDDVYVPGTGNBEGYELMTTDYDNIPANCLNKNE-----YAVGRETRDVF 166  
DB 343 --DSTN-----GDGSNSBEGY---TDYSSISENTYDQIDPNYTYQOTEDTSQYTT 387

RESULT 11  
US-09-107-532A-6865  
Sequence 6865, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Arinello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 6865:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 113 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (B) LOCATION 1...113  
SEQUENCE DESCRIPTION: SEQ ID NO: 6865:  
US-09-107-532A-6865  
Query Match 7.9%; Score 73.5; DB 4; Length 113;  
Best Local Similarity 29.4%; Pred. No. 1.3;  
Matches 32; Conservative 8; Mismatches 54; Indels 15; Gaps 5;  
QY 67 FLFNNADTNMQPATEKTAIVAMGYKRNEMARFETEDGQVTFDVIAYS-----DNDQDV 121  
DB 1 FQVMEVADTMRLRYRFSFNKNDGYELSKTRIRERTINDSYIKETILYSVARLEKNOLI 60  
QY 122 IYVPG--TDGNEBEGYELMTTDYDNIPANCL-NKREYAVGRETRDVF 167  
DB 61 VSYPGEVTHGKRRYRYIT---EIGRLLEKKQEWLT---TDVYTS 102

RESULT 12  
US-09-107-532A-4711  
Sequence 4711, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <unknown>

SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107.532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariadello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 4711:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 489 amino acids  
TYPE: amino acid  
MOLECULE TYPE: linear  
HYPOTHEICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (B) LOCATION 1...489  
SEQUENCE DESCRIPTION: SEQ ID NO: 4711:  
US-09-107-532A-4711

Query Match 7.9%; Score 73.5; DB 4; Length 489;  
Best Local Similarity 23.5%; Pred. No. 11;  
Matches 32; Conservative 19; Mismatches 30; Indels 55; Gaps 8;  
QY 24 DVENVYMKATYKND-----DPWGNDFTCVGMADVDEKSIQAEFLF-----69  
DB 44 DFEETPIVETAYSDLGARPSIDPV-----ILV-----KLVIQYLFGRSMRQT 87  
QY 70 MNNAIDYVWQPA-----TEKVTAVKMYGYNRENAFRETEDGQVFTDVIAY-----114  
DB 88 IKEDFTVAIRWFLGYSFEERKIPHSFGKMYRRFRET-----VFEDIRAYILQAVKA 143  
QY 115 ---SDNCQDVIVPGT 127  
DB 144 GFVTEEDN---LYLDST 156

RESULT 13  
US-09-134-001C-3716  
Sequence 3716, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 3716  
LENGTH: 1335  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3716

Query Match 7.8%; Score 73; DB 4; Length 1335;  
Best Local Similarity 23.7%; Pred. No. 53;  
Matches 36; Conservative 18; Mismatches 46; Indels 52; Gaps 9;

QY 19 KSLKAD-VENY---YMKATYKND---PYWGNDFTCVGMADVDEKSIQAEFLFMNN 72  
DB 329 KMKAPRIEEDYTSYFKYGRNGVGRP-----EGIVYHDTANDNSTIIGELAFMKR 380  
QY 73 ADTNMOPATEKRYTAVKMYGYNRENAFRETEDGQVFTDVIAYSDNCQDVIVPGT 127  
DB 381 -----NYTNAFVHARVDGNRIETAPFTDYLMSG-----AGP 411  
QY 128 DGNIEGYE---LWTTDNDINPACLNKFNEYA 156  
DB 412 YGNORFINVEIVHVDYDSE-ARSMNNYADYA 442  
RESULT 14  
5268463-2  
Patent No. 5268463  
APPLICANT: JEFFERSON, RICHARD A.  
TITLE OF INVENTION: PLANT PROMOTER a-GLUCURONIDASE GENE  
CONSTRUCT  
NUMBER OF SEQUENCES: 9  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/447,976  
FILING DATE: 08-DEC-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 119,102  
FILING DATE: 10-NOV-1987  
APPLICATION NUMBER: 264,586  
FILING DATE: 31-OCT-1988  
SEQ ID NO: 2  
LENGTH: 600  
5268463-2

Query Match 7.7%; Score 72.5; DB 6; Length 600;  
Best Local Similarity 23.4%; Pred. No. 19;  
Matches 41; Conservative 26; Mismatches 55; Indels 53; Gaps 10;  
QY 1 NOPD-WADERANG---AHQDAWKSILKA-----DVE-----NYYVW 32  
DB 368 NKRELYSEAVANGETQOAHLOAKELIADKNHPSVMMISANEDDTROVHGNSPLA 427  
QY 33 KATYKNDPYWGNDFTCVGMADVDEKSIQAEFLFMNNAIDYVWQPA-----RYVGM 92  
DB 428 EATRKIDPT--RPTCVNWFCDARHDTISDLDFVLCN-----RYYGM 469  
QY 93 NRENAFRETEDGQVFTDVIAYSDNCQDVIVP--GTD---GNEEGY-ELWTTDY 141  
DB 470 YVDSG-DLETAERVLEKELIAWQEKLPITITTEYGVDTLAGLSMTDMSSEY 523

RESULT 15  
US-08-882-704A-5  
Sequence 5, Application US/08882704A  
Patent No. 5879906  
GENERAL INFORMATION:  
APPLICANT: Jefferson, Richard A.  
APPLICANT: Wilson, Katherine J.  
APPLICANT: Leader, Michael  
TITLE OF INVENTION: GLUCURONIDE REPRESSORS AND USES THEREOF  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
City: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2003, 12:15:04 : Search time 53.0914 Seconds

(Without alignments)  
511,236 Million cell updates/sec

Title: US-10-087-195-7

Perfect score: 936

Sequence: 1 NQPMWDEANCAHQDAWKS.....FNEYAVGRETRDYFTSACLE 171

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	936	100.0	171	22	AA1980.DAT.*
2	936	100.0	171	22	AA1981.DAT.*
3	936	100.0	190	19	AA1982.DAT.*
4	936	100.0	190	20	AA1983.DAT.*
5	630	67.3	172	22	AA1984.DAT.*
6	630	67.3	172	22	AA1985.DAT.*
7	630	67.3	190	19	AA1986.DAT.*
8	630	67.3	190	20	AA1987.DAT.*
9	331.5	35.4	182	22	AA1988.DAT.*

10	331.5	35.4	182	22	AA1989.DAT.*
11	331.5	35.4	200	19	AA1990.DAT.*
12	331.5	35.4	200	20	AA1991.DAT.*
13	288	30.8	203	20	AA1992.DAT.*
14	285	30.4	203	20	AA1993.DAT.*
15	282	30.1	209	19	AA1994.DAT.*
16	282	30.1	209	20	AA1995.DAT.*
17	222	23.7	198	20	AA1996.DAT.*
18	113	12.1	285	20	AA1997.DAT.*
19	105	11.2	1592	14	AA1998.DAT.*
20	99	10.6	284	20	AA1999.DAT.*
21	91.5	9.8	221	22	AA2000.DAT.*
22	87	9.3	338	22	AA2001.DAT.*
23	86.5	9.2	738	23	AA2002.DAT.*
24	85.5	9.1	337	23	AA2003.DAT.*
25	85.5	9.1	1612	23	AA2004.DAT.*
26	84.5	9.0	408	22	AA2005.DAT.*
27	84	9.0	624	17	AA2006.DAT.*
28	83.5	8.9	249	21	AA2007.DAT.*
29	83.5	8.9	275	21	AA2008.DAT.*
30	83.5	8.9	2911	24	AA2009.DAT.*
31	83.5	8.9	2912	22	AA2010.DAT.*
32	82.5	8.8	405	17	AA2011.DAT.*
33	82.5	8.8	954	21	AA2012.DAT.*
34	82.5	8.8	954	21	AA2013.DAT.*
35	81.5	8.7	1129	22	AA2014.DAT.*
36	80.5	8.6	516	22	AA2015.DAT.*
37	80	8.5	240	21	AA2016.DAT.*
38	80	8.5	249	21	AA2017.DAT.*
39	80	8.5	249	22	AA2018.DAT.*
40	80	8.5	321	20	AA2019.DAT.*
41	80	8.5	929	22	AA2020.DAT.*
42	79	8.4	327	22	AA2021.DAT.*
43	79	8.4	327	22	AA2022.DAT.*
44	78.5	8.4	1752	24	AA2023.DAT.*
45	78.5	8.4	1752	24	AA2024.DAT.*

#### ALIGNMENTS

RESULT 1  
AAB74289 standard; protein: 171 AA.  
AAB74289:  
AC AAB74289:  
XX 20-JUN-2001 (first entry)  
XX 20-JUN-2001 (first entry)  
XX Histacalain protein FS-HBP2.  
XX Histacalain; FS-HBP1, conjunctivitis.  
XX Unidentified.  
XX WO200115719-A2.  
XX 08-MAR-2001.  
XX 24-AUG-2000; 2000WO-GB03282.  
XX 01-SEP-1999; 99GB-0020674.  
XX (EVOL-) EVOLUTEC LTD.  
XX Nuttall PA, Paesen GC;  
XX WPI, 2001-257675/26.  
XX Use of histacalain proteins for treating or preventing non-infective  
XX conjunctivitis, or for manufacturing a medicament for treating or  
XX preventing conjunctivitis, e.g. seasonal or perennial allergic  
XX conjunctivitis

XX Claim 4; Page 5-6; 19pp; English.  
PS  
CC The present invention relates to the use of a histacalin protein  
CC for treating or preventing conjunctivitis. The present sequence  
CC is the histacalin protein-FS-HBP2. The invention is particularly useful  
CC in the treatment of allergic or seasonal conjunctivitis.  
XX  
SQ Sequence 171 AA;  
Query Match 100.0%; Score 936; DB 22; Length 171;  
Best Local Similarity 100.0%; Pred. No. 1.6e-89;  
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NOPDMADEAANGAHODAMKSLKADVENYVMKATYKNDPVWGNDFTCVGMANDVNEDE 60  
DB 1 NOPDMADEAANGAHODAMKSLKADVENYVMKATYKNDPVWGNDFTCVGMANDVNEDE 60  
QY 61 KSIQAEFLFNNNDTNMQFATEKYTAVKMGYNRENAFRYETEDGQVFTVIAVSDNCD 120  
DB 61 KSIQAEFLFNNNDTNMQFATEKYTAVKMGYNRENAFRYETEDGQVFTVIAVSDNCD 120  
QY 121 VIIVPGTGDNEGEYELMTTDDYDNIIPANCLKFNFEYAVGRTROYFTSACLE 171  
DB 121 VIIVPGTGDNEGEYELMTTDDYDNIIPANCLKFNFEYAVGRTROYFTSACLE 171  
RESULT 2  
AAB73261  
ID AAB73261 standard; protein; 171 AA.  
AC AAB73261;  
DT 15-MAY-2001 (first entry)  
DE Histacalin protein FS-HBP2.  
XX  
XX FS-HBP2; histacalin protein; antiinflammatory; antiallergic;  
KW ophthalmological; allergic rhinitis.  
XX  
XX Unidentified.  
XX  
XX WO200116164-A2.  
XX  
XX PD 08-MAR-2001.  
XX  
XX PF 24-AUG-2000; 2000WO-GB03287.  
XX  
XX PR 01-SEP-1999; 99GB-0020673.  
XX  
XX (EVOL-) EVOLUTEC LTD.  
XX  
XX Nuttall PA, Paesen GC;  
XX  
XX WPI; 2001-218521/22.  
XX  
XX PT Use of histacalin proteins for treating or preventing allergic  
PT rhinitis, or for manufacturing a medicament for treating or preventing  
PT allergic rhinitis, e.g. seasonal or perennial allergic rhinitis  
XX  
XX PS Disclosure; Pages 4-6; 19pp; English.  
XX  
XX CC The present invention relates to a method for treating or preventing  
CC allergic rhinitis. The method involves employing a blood-feeding  
CC ectoparasite-derived (e.g. tick-derived) histacalin protein. The present  
CC sequence is one such histacalin protein. The histacalin protein, is  
CC useful for treating or preventing allergic rhinitis, both seasonal and  
CC perennial allergic conjunctivitis.  
XX  
SQ Sequence 171 AA;  
Query Match 100.0%; Score 936; DB 22; Length 171;  
Best Local Similarity 100.0%; Pred. No. 1.6e-89;

Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NOPDMADEAANGAHODAMKSLKADVENYVMKATYKNDPVWGNDFTCVGMANDVNEDE 60  
DB 1 NOPDMADEAANGAHODAMKSLKADVENYVMKATYKNDPVWGNDFTCVGMANDVNEDE 60  
QY 61 KSIQAEFLFNNNDTNMQFATEKYTAVKMGYNRENAFRYETEDGQVFTVIAVSDNCD 120  
DB 61 KSIQAEFLFNNNDTNMQFATEKYTAVKMGYNRENAFRYETEDGQVFTVIAVSDNCD 120  
QY 121 VIIVPGTGDNEGEYELMTTDDYDNIIPANCLKFNFEYAVGRTROYFTSACLE 171  
DB 121 VIIVPGTGDNEGEYELMTTDDYDNIIPANCLKFNFEYAVGRTROYFTSACLE 171  
RESULT 3  
AAW37447  
ID AAW37447 standard; protein; 190 AA.  
AC AAW37447;  
DT 08-JUN-1998 (first entry)  
DE Tick vasoactive amine binding protein 2 FS-HBP2.  
XX  
XX Female-specific vasoactive amine binding protein 1; FS-HCP1;  
KW histamine; serotonin; assay; antihistamine; anti-inflammatory;  
KW insect bite; snake bite; scorpion bite; dermatitis; vaccine;  
KW transgenic animal; tick.  
XX  
XX OS Rhipicephalus appendiculatus.  
XX  
XX FH Key Location/Qualifiers  
FT Peptide 1..19  
FT /label= Sig\_peptide  
XX  
XX WO9744451-A2.  
XX  
XX PD 27-NOV-1997.  
XX  
XX PF 19-MAY-1997; 97WO-GB01372.  
XX  
XX PR 18-APR-1997; 97GB-0007844.  
XX  
XX PR 18-MAY-1996; 96GB-0010484.  
XX  
XX PA (OXFO-) OXFORD VACS LTD.  
XX  
XX PI Nuttall PA, Paesen GC;  
XX  
XX WPI; 1998-018506/02.  
XX  
XX DR N-PSDB; AAW00228.  
XX  
XX PT New vasoactive amine binding proteins and related nucleic acid,  
PT vectors - transformed cells and transgenic animals, used for  
PT assaying or removing histamine and as antihistamine or  
PT anti-inflammatory agents  
XX  
XX PS Example 2; Fig 2; 44pp; English.  
XX  
XX CC This protein comprises tick Rhipicephalus appendiculatus (Ra) novel  
CC female-specific histamine binding protein 2 (FS-HBP2). Its amino  
CC acid sequence was deduced from a cDNA clone (see AAW00228) obtained  
CC from a salivary gland cDNA library. FS-HBP1 (see AAW37446) and  
CC male-specific HSP1 (see AAW37448) and a related protein, D-RET6 (see  
CC AAW37449) from Dermacentor reticulatus, were also identified. These  
CC novel vasoactive amine binding proteins (VABPs) can be expressed  
CC in host cells using e.g. a baculovirus expression system. They can  
CC be used: (i) to assay histamine (or other VA such as serotonin) in  
CC body fluids or cell culture supernatants, e.g. to monitor the  
CC effect of allergens; (ii) for binding VA, e.g. to remove histamine  
CC from blood, food, cell cultures etc.; (iii) as an antihistamine or  
CC anti-inflammatory agents, e.g. for treating insect, snake or  
CC scorpion bites or dermatitis, or as a carrier for slow release of

CC histamine-related compounds; (iv) in vaccines to protect against  
CC metazoan parasites, especially in animals; (v) as reagents for  
CC studying inflammation, involvement of VA in ulcer formation or the  
CC immune response etc. VABPs provide a more sensitive assay for  
CC histamine than low affinity antibodies currently used. They may  
CC also be more effective and safer than conventional antihistamines.

SQ Sequence 190 AA;

Query Match 100.0%; Score 936; DB 19; Length 190;  
Best Local Similarity 100.0%; Pred. No. 1.8e-89;  
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NOPDMADEAANGAHODAMKSLKADVENYVYWKATYKKNDPVGNGFTCGVANDVNEDE 60  
DB 20 NOPDMADEAANGAHODAMKSLKADVENYVYWKATYKKNDPVGNGFTCGVANDVNEDE 79  
OY 61 KSIQAEFLFMNADTNMOPATEKYAVKMYGYNRENAFRYETEDGQVFTDVIAYSDDNCD 120  
DB 80 KSIQAEFLFMNADTNMOPATEKYAVKMYGYNRENAFRYETEDGQVFTDVIAYSDDNCD 139  
OY 121 VIYVGTGNGEGYELMTTDDNIPANCLNKFNEAVGERETRDVFTSACLE 171  
DB 140 VIYVGTGNGEGYELMTTDDNIPANCLNKFNEAVGERETRDVFTSACLE 190

#### RESULT 4

AAV18079 standard; Protein: 190 AA.

AAV18079;

06-AUG-1999 (first entry)

Histamine binding protein FS-HBP2.

KW Histamine binding protein; serotonin binding compound; inflammation;  
KW gastric acid secretion; allergy; type I hypersensitivity reaction;  
KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
KW respiratory disease; coronary heart disease; cellular growth regulator;  
KW tissue repair; blood-sucking ectoparasite; therapy.

Rhipicephalus appendiculatus.

W09927104-A1.

03-JUN-1999.

26-NOV-1998; 98MO-GB03530.

26-JUN-1998; 98GB-0013917.

26-NOV-1997; 97GB-0025046.

(OXFO-) OXFORD VACS LTD.

Nuttall PA, Paesen GC;

WPI: 1999-357841/30.

N-PSDB: AAX76965.

Histamine and serotonin binding compounds useful for the treatment

of allergies

Claim 1; Fig 2; 84pp; English.

CC This sequence is an example of a histamine or serotonin binding  
CC compound (A), of the invention. The compounds are useful for regulating  
CC the action of histamine and serotonin (in e.g. inflammation and gastric  
CC acid secretion), the detection, quantification and removal of histamine  
CC or serotonin (in animals, plants, cell cultures, food materials, or  
CC humans) and in the treatment of various diseases and allergies  
CC (e.g. type I hypersensitivity reactions, urticaria, asthma, allergic

CC rhinitis (hay fever), atopic dermatitis, insect bites and food and drug  
CC allergies, abnormal blood pressure, migraine, psychological disorders,  
CC respiratory disease, and coronary heart disease). Histamine may also be  
CC used to regulate cellular growth and tissue repair. The molecules may  
CC also be used as components of vaccines directed against blood-sucking  
CC ectoparasites.

SQ Sequence 190 AA;

Query Match 100.0%; Score 936; DB 20; Length 190;  
Best Local Similarity 100.0%; Pred. No. 1.8e-89;  
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NOPDMADEAANGAHODAMKSLKADVENYVYWKATYKKNDPVGNGFTCGVANDVNEDE 60  
DB 20 NOPDMADEAANGAHODAMKSLKADVENYVYWKATYKKNDPVGNGFTCGVANDVNEDE 79  
OY 61 KSIQAEFLFMNADTNMOPATEKYAVKMYGYNRENAFRYETEDGQVFTDVIAYSDDNCD 120  
DB 80 KSIQAEFLFMNADTNMOPATEKYAVKMYGYNRENAFRYETEDGQVFTDVIAYSDDNCD 139  
OY 121 VIYVGTGNGEGYELMTTDDNIPANCLNKFNEAVGERETRDVFTSACLE 171  
DB 140 VIYVGTGNGEGYELMTTDDNIPANCLNKFNEAVGERETRDVFTSACLE 190

#### RESULT 5

AAB74288 standard; Protein: 172 AA.

AAB74288;

20-JUN-2001 (first entry)

Histacalin protein FS-HBPL.

KW Histacalin; FS-HBPL; conjunctivitis.  
KW unidentified.  
KW W0200115719-A2.

08-MAR-2001.

24-AUG-2000; 2000MO-GB03282.

01-SEP-1999; 99GB-0020674.

(EVOL-) EVOLUTEC LTD.

Nuttall PA, Paesen GC;

WPI: 2001-257675/26.

Use of histacalin proteins for treating or preventing non-infective

conjunctivitis, or for manufacturing a medicament for treating or

preventing conjunctivitis, e.g. seasonal or perennial allergic

conjunctivitis

Claim 4; page 5-6; 19pp; English.

The present invention relates to the use of a histacalin protein

for treating or preventing conjunctivitis. The present sequence

is the histacalin protein FS-HBPL. The invention is particularly useful

in the treatment of allergic or seasonal conjunctivitis.

SQ Sequence 172 AA;

Query Match 67.3%; Score 630; DB 22; Length 172;  
Best Local Similarity 66.5%; Pred. No. 1.3e-57;  
Matches 113; Conservative 20; Mismatches 37; Indels 0; Gaps 0;

OY 1 NOPDMADEAANGAHODAMKSLKADVENYVYWKATYKKNDPVGNGFTCGVANDVNEDE 60

QY 121 VIYVPGTSGNEEGYELWTIDYDINIPANCLKNFNEYAVGRETRDVFYSACL 1700  
||| | ||: |||| ||: |||| |||| | ||: || ||  
Db 121 VYIALGPDSGAGYELMAFDYTDVPASCLEKFENEYAAGLPIYRDVYTSACL 1700

[illegible]

```

Query Match      67.3%; Score 630; DB 19; Length 190;
Best Local Similarity 66.5%; Pred. No. 1.5e-57;
Matches 113; Conservative 20; Mismatches 37; Indels 0; Gaps 0
QY      1 NQPDADAEANAHDDAKMSLKADYENYYVVKATYTNDDYVWGNDFTCYGVMAINDVNEDE 60
:::|||||  ||||| : || : :|||  |||||  ||| : |||

```

Db 19 DKPWADDEANGEHQDAMKHLQKLVENYDLIKATYKNDPVGNDFTCVGTAAQNLNDE 78  
Qy 61 KSTQAEFLFMNNADTNNGFATEKVTAKMGYNRENAFRYETEDGCVFTVIAVSDNCD 120  
Db 79 KNEAEMFMNNADTVYQHTFEKATPKMGYNENATYOTEDGCVLTVLAVSDNDCY 138  
Qy 121 VIIVPGTGDGNEGEYELMTDTYDNIIPANCLNKFNEYAVGRETVDFTSACL 170  
Db 139 VITALGPDGSGAGIELMATDTYTPASCLEKFNERYAAGLVPRDYVTSACL 188

RESULT 8  
AAV18078  
ID AAV18078 standard; Protein: 190 AA.  
XX  
AC AAV18078;  
XX  
DT 06-AUG-1999 (first entry)  
XX  
DE Histamine binding protein FS-HBPI.  
XX  
KW Histamine binding protein; serotonin binding compound; inflammation;  
KW gastric acid secretion; allergy; type I hypersensitivity reaction;  
KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
KW respiratory disease; coronary heart disease; cellular growth regulator;  
KW tissue repair; blood-sucking ectoparasite; therapy.  
XX  
OS Rhinopcephalus appendiculatus.  
XX  
PN MO9927104-A1.  
XX  
PD 03-JUN-1999.  
XX  
PE 26-NOV-1998; 98WO-G803530.  
XX  
PR 26-JUN-1998; 98GB-0013917.  
XX  
PR 26-NOV-1997; 97GB-0025046.  
XX  
PA (OXFO-) OXFORD VACS LTD.  
XX  
PI Nuttall PA, Paesen GC;  
XX  
DR WPI: 1999-357841/30.  
XX  
DR N-PSDB: AAX76964.  
XX  
PT Histamine and serotonin binding compounds useful for the treatment  
PT of allergies  
XX  
PS Claim 1; Fig 1; 84pp; English.  
XX  
CC This sequence is an example of a histamine or serotonin binding  
CC compound (A). Of the invention. The compounds are useful for regulating  
CC the action of histamine and serotonin (in e.g. inflammation and gastric  
CC acid secretion), the detection, quantification and removal of histamine  
CC or serotonin (in animals, plants, cell cultures, food materials, or  
CC humans) and in the treatment of various diseases and allergies  
CC (e.g. type I hypersensitivity reactions, urticaria, asthma, allergic  
CC rhinitis (hay fever), atopic dermatitis, insect bites and food and drug  
CC allergies, abnormal blood pressure, migraine, psychological disorders,  
CC respiratory disease, and coronary heart disease). Histamine may also be  
CC used to regulate cellular growth and tissue repair. The molecules may  
CC also be used as components of vaccines directed against blood-sucking  
CC ectoparasites.  
XX  
SQ Sequence 190 AA:  
Query Match 67.3%; Score 630; DB 20: length 190;  
Best Local Similarity 66.5%; Pred. No. 1.5e-57;  
Matches 113; Conservative 20; Mismatches 37; Indels 0; Gaps 0;  
Qy 1 NOPWADDEANGAHQDAMKSLKADVENYVWKATYKNDPVGNDFTCVGVMANDVNEDE 60  
::: ||||| ||||| : | : ||||| ||||| ||||| : |||||

Db 19 DKPWADDEANGEHQDAMKHLQKLVENYDLIKATYKNDPVGNDFTCVGTAAQNLNDE 78  
Qy 61 KSTQAEFLFMNNADTNNGFATEKVTAKMGYNRENAFRYETEDGCVFTVIAVSDNCD 120  
Db 79 KNEAEMFMNNADTVYQHTFEKATPKMGYNENATYOTEDGCVLTVLAVSDNDCY 138  
Qy 121 VIIVPGTGDGNEGEYELMTDTYDNIIPANCLNKFNEYAVGRETVDFTSACL 170  
Db 139 VITALGPDGSGAGIELMATDTYTPASCLEKFNERYAAGLVPRDYVTSACL 188

RESULT 9  
AAB74290  
ID AAB74290 standard; protein: 182 AA.  
XX  
AC AAB74290;  
XX  
DT 20-JUN-2001 (first entry)  
XX  
DE Histacalin protein MS-HBPI.  
XX  
KW Histacalin; MS-HBPI; conjunctivitis.  
XX  
OS Unidentified.  
XX  
PN WO200115719-A2.  
XX  
PD 08-MAR-2001.  
XX  
PE 24-AUG-2000; 2000WO-G803282.  
XX  
PR 01-SEP-1999; 99GB-0020674.  
XX  
PA (EVOL-) EVOLUTEC LTD.  
XX  
PI Nuttall PA, Paesen GC;  
XX  
DR WPI: 2001-257675/26.  
XX  
PT Use of histacalin proteins for treating or preventing non-infective  
PT conjunctivitis, or for manufacturing a medicament for treating or  
PT preventing conjunctivitis, e.g. seasonal or perennial allergic  
PT conjunctivitis  
XX  
PS Claim 4; Page 5-6; 19pp; English.  
XX  
CC The present invention relates to the use of a histacalin protein  
CC for treating or preventing conjunctivitis. The present sequence  
CC is the histacalin protein FS-HBPI. The invention is particularly useful  
CC in the treatment of allergic or seasonal conjunctivitis.  
XX  
SQ Sequence 182 AA:  
Query Match 35.4%; Score 331.5; DB 22: length 182;  
Best Local Similarity 42.0%; Pred. No. 2e-26;  
Matches 74; Conservative 29; Mismatches 62; Indels 11; Gaps 7;  
Qy 3 PDWADDEANGAHQDAMKSLKADVENYVWKATYKNDPVGNDFTCVGVMANDVNEDEKS 62  
Db 2 PTWANEKLSYODAMSLQDOKRYYLAOTDTTGWGEPTCVSVAEKIG--KKK 59  
Qy 63 IOAEFLFMNNADTNNGFATEKVTAKMGYNRENAFRYETED--GOVFTVIAVSD--DNC 119  
Db 60 LNAITILYKNNKHLIDKASHETITVWKAVIDTTEGIGITETOGTTFEDVFPVSDYKNC 119  
Qy 120 DVIVPGTGDGNEGEYELMTDTD--YDNIIPANCLNKFNEYAVGRE--TRDVFT--SAC 169  
Db 120 DVIVPKEKRSDEDEYELAWSEDKIKIPDCKFTMAVFAQQGKRYRNVYTDSAC 175

RESULT 10  
AAB73262  
ID AAB73262 standard; protein: 182 AA.

```

XX AAB73262;
XX
XX 15-MAY-2001 (first entry)
XX
XX Histacalin protein MS-HBPI.
XX
XX MS-HBPI; histacalin protein; antiinflammatory; antiallergic;
XX ophthalmological; allergic rhinitis.
XX
XX Unidentified.
XX
XX W0200116164-A2.
XX
XX 08-MAR-2001.
XX
XX 24-AUG-2000; 2000WO-GB03287.
XX
XX 01-SEP-1999; 99GB-0020673.
XX
XX (EVOL-) EVOLUTEC LTD.
XX
XX Nuttall PA, Paesen GC;
XX
XX WPI; 2001-218521/22.
XX
XX Use of histacalin proteins for treating or preventing allergic
XX rhinitis, or for manufacturing a medicament for treating or preventing
XX allergic rhinitis, e.g. seasonal or perennial allergic rhinitis -
XX
XX Disclosure; Pages 4-6; 19pp; English.
XX
XX The present invention relates to a method for treating or preventing
XX allergic rhinitis. The method involves employing a blood-feeding
XX ectoparasite-derived (e.g. tick-derived) histacalin protein. The present
XX sequence is one such histacalin protein. The histacalin protein, is
XX useful for treating or preventing allergic rhinitis, both seasonal and
XX perennial allergic conjunctivitis.
XX
XX Sequence 182 AA;
XX
XX Query Match 35.4%; Score 331.5; DB 22; Length 182;
XX Best Local Similarity 42.0%; Pred. No. 2e-26;
XX Matches 74; Conservative 29; Mismatches 62; Indels 11; Gaps 7;
XX
XX 3 PMWADAEANGAHODAMKSLKADYENYYWAKATYKNDPVGNDFTCVGVANDVNDDEKS 62
XX 2 PTWANEAKLGSYODAKMSLQODONKRYTLAQATQTDGVGGEFTCVSYAEKIG--KKK 59
XX
XX 63 IOAEFLPMNADTNMOPATEKVTAVKMGYNRENARFETED--GGVFTDVIAYS-DNC 119
XX 60 LNAFTILYKKNHLLDLKESHETITVWKAIDYTTENGKIKYETQGRTOFFEDVFEFSDYKNC 119
XX
XX 120 DVIYVPGTDGNEG-YELWTTD--YDNIPANCINKREYAVGRE--TRDVF--SAC 169
XX 120 DVIYVPERKSGDEGDELWVSEDKIDPCCKFTMAYFAQOQEKTVRVNYTDS SC 175
XX
XX RESULT 11
XX AAM37448 standard; Protein; 200 AA.
XX
XX AAM37448;
XX
XX 08-JUN-1998 (first entry)
XX
XX Tick vasoactive amine binding protein 1 MS-HBPI.
XX
XX Male-specific vasoactive amine binding protein 1; MS-HBPI;
XX histamine; serotonin; assay: antihistamine; anti-inflammatory;
XX insect bite; snake bite; scorpion bite; dermatitis; vaccine;
XX transgenic animal; tick.
XX

```

```

OS Rhipicephalus appendiculatus.
XX
XX Key Location/Qualifiers
XX Peptide 1..18
XX Modified-site /label= Sig-peptide
XX 79..81
XX /note= "Asn is N-glycosylated"
XX
XX W09744451-A2.
XX
XX 27-NOV-1997.
XX
XX 19-MAY-1997; 97WO-GB01372.
XX
XX 18-APR-1997; 97GB-0007844.
XX 18-MAY-1996; 96GB-0010484.
XX
XX (OXFO-) OXFORD VACS LTD.
XX
XX Nuttall PA, Paesen GC;
XX
XX WPI; 1998-018506/02.
XX N-PSDB; AAV00229.
XX
XX New vasoactive amine binding proteins and related nucleic acid,
XX vectors - transformed cells and transgenic animals, used for
XX assaying or removing histamine and as antihistamine or
XX anti-inflammatory agents
XX
XX Example 2; Fig 3; 44pp; English.
XX
XX This protein comprises tick Rhipicephalus appendiculatus (Ra) novel
XX male-specific histamine binding protein 1 (MS-HBPI). Its amino
XX acid sequence was deduced from a cDNA clone (see AAV00227) obtained
XX from a salivary gland cDNA library. Female-specific HSP1 and
XX HSP2 (see AAM37446-47) and a related protein, D.RET6 (see AAM37449)
XX from Dermacentor reticulatus, were also identified. These novel
XX vasoactive amine binding proteins (VABPs) can be expressed in
XX host cells using e.g. a baculovirus expression system. They can
XX be used: (i) to assay histamine (or other VA such as serotonin) in
XX body fluids or cell culture supernatants, e.g. to monitor the
XX effect of allergens; (ii) for binding VA, e.g. to remove histamine
XX from blood, food, cell cultures etc.; (iii) as an antihistamine or
XX anti-inflammatory agents, e.g. for treating insect, snake or
XX scorpion bites or dermatitis; or as a carrier for slow release of
XX histamine-related compounds; (iv) in vaccines to protect against
XX metazoan parasites, especially in animals; (v) as reagents for
XX studying inflammation. Involvement of VA in ulcer formation or the
XX immune response etc. VABPs provide a more sensitive assay for
XX histamine than low-affinity antibodies currently used. They may
XX also be more effective and safer than conventional antihistamines.
XX
XX Sequence 200 AA;
XX
XX Query Match 35.4%; Score 331.5; DB 19; Length 200;
XX Best Local Similarity 42.0%; Pred. No. 2.3e-26;
XX Matches 74; Conservative 29; Mismatches 62; Indels 11; Gaps 7;
XX
XX 3 PMWADAEANGAHODAMKSLKADYENYYWAKATYKNDPVGNDFTCVGVANDVNDDEKS 62
XX 20 PTWANEAKLGSYODAKMSLQODONKRYTLAQATQTDGVGGEFTCVSYAEKIG--KKK 77
XX
XX 63 IOAEFLPMNADTNMOPATEKVTAVKMGYNRENARFETED--GGVFTDVIAYS-DNC 119
XX 78 LNAFTILYKKNHLLDLKESHETITVWKAIDYTTENGKIKYETQGRTOFFEDVFEFSDYKNC 137
XX
XX 120 DVIYVPGTDGNEG-YELWTTD--YDNIPANCINKREYAVGRE--TRDVF--SAC 169
XX 138 DVIYVPERKSGDEGDELWVSEDKIDPCCKFTMAYFAQOQEKTVRVNYTDS SC 193
XX
XX RESULT 12
XX AAY18080
XX

```

ID AAY18080 standard; Protein: 200 AA.  
XX  
AC AAY18080;  
XX  
DT 06-AUG-1999 (first entry)  
XX  
DE Histamine binding protein MS-HBPL.  
XX  
KW Histamine binding protein; serotonin binding compound; inflammation;  
KW gastric acid secretion; allergy; type I hypersensitivity reaction;  
KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
KW respiratory disease; coronary heart disease; cellular growth regulator;  
KW tissue repair; blood-sucking ectoparasite; therapy.  
XX  
OS Bhocephalus appendiculatus.  
XX  
PN WO9927104-A1.  
XX  
PD 03-JUN-1999.  
XX  
PF 26-NOV-1998; 98WO-GB03530.  
XX  
PR 26-JUN-1998; 98GB-0013917.  
XX  
PR 26-NOV-1997; 97GB-0025046.  
XX  
PA (OXFO-) OXFORD VACS LTD.  
XX  
PI Nuttall PA, Paesen GC;  
XX  
DR WPI: 1999-357841/30.  
XX  
DR N-PSDB: AAX76966.  
XX  
PT Histamine and serotonin binding compounds useful for the treatment  
XX  
PT of allergies  
XX  
PS Claim 1; Fig 3; 84pp; English.  
XX  
CC This sequence is an example of a histamine or serotonin binding  
XX compound (A), of the invention. The compounds are useful for regulating  
XX the action of histamine and serotonin (in e.g. inflammation and gastric  
XX acid secretion), the detection, quantification and removal of histamine  
XX or serotonin (in animals, plants, cell cultures, food materials, or  
XX humans) and in the treatment of various diseases and allergies  
XX (e.g. type I hypersensitivity reactions, urticaria, asthma, allergic  
XX rhinitis (hay fever), atopic dermatitis, insect bites and food and drug  
XX allergies, abnormal blood pressure, migraine, psychological disorders,  
XX respiratory disease, and coronary heart disease). Histamine may also be  
XX used to regulate cellular growth and tissue repair. The molecules may  
XX also be used as components of vaccines directed against blood-sucking  
XX ectoparasites.  
XX  
SQ Sequence 200 AA:  
XX  
Query Match 35.4%; Score 331.5; DB 20; Length 200;  
Best Local Similarity 42.0%; Pred. No. 2.3e-26;  
Matches 74; Conservative 29; Mismatches 62; Indels 11; Gaps 7;  
OY 3 PDMADEANGAHODAMKSLKADVENYVMKATYKNDPVNGNDFTCGVMANDVNEDEKS 62  
DB 20 PTMADEANGAHODAMKSLKADVENYVMKATYKNDPVNGNDFTCGVMANDVNEDEKS 77  
OY 63 IOAERLFPMNADTNMOPATEKYAVKMGYRENAFRETEDE--GQVTDVIANSD-DNC 119  
DB 78 LNAITLYKNKHLTDLKESHERITWKADYTTENGIKETGTROTDEDFEVESDYDNC 137  
OY 120 DVIVPPTGDNBEG-YELMTD--YDNIPANCLKNFENYAVGRE--TRDVPF-SAC 169  
DB 138 DVIVPPTGDNBEG-YELMTD--YDNIPANCLKNFENYAVGRE--TRDVPF-SAC 193  
RESULT 13  
AAY18085

ID AAY18085 standard; Protein: 203 AA.  
XX  
AC AAY18085;  
XX  
DT 06-AUG-1999 (first entry)  
XX  
DE Histamine binding protein 1h/Bm-HBPL.  
XX  
KW Histamine binding protein; serotonin binding compound; inflammation;  
KW gastric acid secretion; allergy; type I hypersensitivity reaction;  
KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
KW respiratory disease; coronary heart disease; cellular growth regulator;  
KW tissue repair; blood-sucking ectoparasite; therapy.  
XX  
OS Bophillus microplus.  
XX  
OS Ixodes hexagonus.  
XX  
OS Synthetic.  
XX  
PN WO9927104-A1.  
XX  
PD 03-JUN-1999.  
XX  
PF 26-NOV-1998; 98WO-GB03530.  
XX  
PR 26-JUN-1998; 98GB-0013917.  
XX  
PR 26-NOV-1997; 97GB-0025046.  
XX  
PA (OXFO-) OXFORD VACS LTD.  
XX  
PI Nuttall PA, Paesen GC;  
XX  
DR WPI: 1999-357841/30.  
XX  
DR N-PSDB: AAX76970.  
XX  
PT Histamine and serotonin binding compounds useful for the treatment  
XX  
PT of allergies  
XX  
PS Claim 13; Fig 7; 84pp; English.  
XX  
CC This sequence is an example of a histamine or serotonin binding  
XX compound (A), of the invention. cDNA encoding this sequence was isolated  
XX from a mixed Bophillus microplus/Ixodes hexagonus cDNA expression  
XX library. The compounds are useful for regulating the action of histamine  
XX and serotonin (in e.g. inflammation and gastric acid secretion), the  
XX detection, quantification and removal of histamine or serotonin (in  
XX animals, plants, cell cultures, food materials, or humans) and in the  
XX treatment of various diseases and allergies (e.g. type I hypersensitivity  
XX reactions, urticaria, asthma, allergic rhinitis (hay fever), atopic  
XX dermatitis, insect bites and food and drug allergies, abnormal blood  
XX pressure, migraine, psychological disorders, respiratory disease, and  
XX coronary heart disease). Histamine may also be used to regulate cellular  
XX growth and tissue repair. The molecules may also be used as components of  
XX vaccines directed against blood-sucking ectoparasites.  
XX  
SQ Sequence 203 AA:  
XX  
Query Match 30.8%; Score 288; DB 20; Length 203;  
Best Local Similarity 35.4%; Pred. No. 8.2e-22;  
Matches 63; Conservative 35; Mismatches 66; Indels 14; Gaps 8;  
OY 1 NOP-----DW-ADEANGAHODAMKSLKADVENYVMKATYKNDPVNGNDFTCGVMAND 55  
DB 27 NEPLKNTWHSKELKN--YDAMWSINOVSTTYTFLSTYNNDSVMSKNTFCLSVTVTS 84  
OY 56 VNEDEKSIOAEFLPMNADTNMOPATEKYAVKMGYRENAFRETEDEGQVTDVIANSD 115  
DB 85 KHESTFVTEYNTTKKNSQOQVSM-TEYNTAVOEGCYVKNKNIQTTENNKTKFNDTVFT 143  
OY 116 D-DNCQDVIVPPTGDNBEGYELMT--TDY-DNIPANCLKNFENYAVGRETRDVPF-SAC 170  
DB 144 DGQTCDLIYIP-----YKENGVELMWRSDYLQNTPTCCQIFIDLVALGRTTYNISPPDCV 198





```

OY      3 PDMADEANGAHODAMKSLKADYENYYMKATYKNDP-VKGNDETCGVMANOVNEDEK 61
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      31 PLMAHEELGKIDAMKSIDGVSVTYVLAKTTIENDTGSWSQFCLQVOEIERKEEDY 90
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      62 SIOAEFLFMNADTNMQF--ATERVTAVKMYGY-NRENAPRYETEDGQVFTDVIAYS-D 117
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      91 TVTSVTEF-RNASSPIKYNYVTEYKAVFQGYKINIRNAIEYQVGGSLNITDILFTDGE 149
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      118 NCDVYIVPGTGDNEGEYELMT--TDYDNIPANCLINKFNEY-AVGRETRDVFYSACL 170
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      150 LCDVFYVPMAD--QGCELWVKKSHYKHVPDYCTFVFNVCACAKDRKTYDIFNEECV 202
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: August 1, 2003, 12:22:48  
 Job time : 55.0914 secs

**THIS PAGE BLANK (USPTO)**



QY 39 LGKYODAMKSIDOGVSIVYIAKTTTENDTSGNSQKCL-----QVOIEKEDDYT 91  
 ID 033983 PRELIMINARY; PRT; 917 AA.  
 AC 033983;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE AMI.  
 GN Listeria monocytogenes.  
 OS Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 OX NCBI\_TaxID=1639;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=EGD;  
 RX MEDLINE=97426036; PubMed=9282740;  
 RA Braun L., Dramsi S., Dehoux P., Bienne H., Lindhal G., Cossart P.;  
 RT "Inlb : an invasion protein of Listeria monocytogenes with a novel  
 type of surface association.";  
 RL Mol. Microbiol. 25:285-294(1997).  
 DR EMBL; U082488; AAC45605.1; -  
 DR Interpro; IPR002502; Amidase\_2.  
 DR Pfam; PF01510; Amidase\_2; 1.  
 DR SMART; SM00644; Ami\_2; 1.  
 SQ SEQUENCE 917 AA; 102352 MW; 21EA2452B4F103B7 CRC64;

## RESULT 12

QY 92 VTSVETFRNASSPIKYYNTETKAVFOYQKNIIRNAIEYOVGGGLN 138  
 ID 033983 PRELIMINARY; PRT; 917 AA.  
 AC 033983;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE AMI.  
 GN Listeria monocytogenes.  
 OS Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 OX NCBI\_TaxID=1639;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=EGD;  
 RX MEDLINE=97426036; PubMed=9282740;  
 RA Braun L., Dramsi S., Dehoux P., Bienne H., Lindhal G., Cossart P.;  
 RT "Inlb : an invasion protein of Listeria monocytogenes with a novel  
 type of surface association.";  
 RL Mol. Microbiol. 25:285-294(1997).  
 DR EMBL; U082488; AAC45605.1; -  
 DR Interpro; IPR002502; Amidase\_2.  
 DR Pfam; PF01510; Amidase\_2; 1.  
 DR SMART; SM00644; Ami\_2; 1.  
 SQ SEQUENCE 917 AA; 102352 MW; 21EA2452B4F103B7 CRC64;

Query Match 7.6%; Score 85.5; DB 2; Length 917;  
 Best Local Similarity 23.2%; Pred. No. 66;  
 Matches 38; Conservative 24; Mismatches 49; Indels 53; Gaps 9;  
 QY 28 GENPLAHEELGKYODAMKSIDOGVSIVYIAKTTTENDTSGNSQKCL-----QVOIEKEDDYT 91  
 ID 033983 PRELIMINARY; PRT; 917 AA.  
 AC 033983;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE AMI.  
 GN Listeria monocytogenes.  
 OS Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 OX NCBI\_TaxID=1639;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=EGD;  
 RX MEDLINE=97426036; PubMed=9282740;  
 RA Braun L., Dramsi S., Dehoux P., Bienne H., Lindhal G., Cossart P.;  
 RT "Inlb : an invasion protein of Listeria monocytogenes with a novel  
 type of surface association.";  
 RL Mol. Microbiol. 25:285-294(1997).  
 DR EMBL; U082488; AAC45605.1; -  
 DR Interpro; IPR002502; Amidase\_2.  
 DR Pfam; PF01510; Amidase\_2; 1.  
 DR SMART; SM00644; Ami\_2; 1.  
 SQ SEQUENCE 917 AA; 102352 MW; 21EA2452B4F103B7 CRC64;

## RESULT 13

QY 325 REAKTSGTITWYQFVSVG-----KITGWDSKALNFTPTPSMEK 363  
 ID 081C77 PRELIMINARY; PRT; 6761 AA.  
 AC 081C77;  
 DT 01-MAR-2003 (TREMBLrel. 23, Created)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE SET-domain protein, putative.  
 GN MAL6P1.131.  
 OS Plasmodium falciparum (isolate 3D7).  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=36329;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC CHEVREACH I., DAVIS P., GOODHEAD I., STEVENS K., MUNGALL K.,  
 RA BERTMAN M., PAIN A., HALL N., ATKIN R., CHILLINGWORTH C., DOGGETT J.,  
 RA ORMOND D., SANDERS M., HAYES R., HALL S., QUAIL M., BARRELL B.;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL844505; CAD50567.1; -  
 SQ SEQUENCE 6761 AA; 797061 MW; 7000C644C420A246 CRC64;

Query Match 7.6%; Score 85.5; DB 5; Length 6761;  
 Best Local Similarity 16.8%; Pred. No. 8.9e+02;  
 Matches 35; Conservative 33; Mismatches 63; Indels 77; Gaps 7;

QY 64 YENDTSGNSQKCLQVOIEKEDDYTVSVETFRNASSPIKYYNTETKAVFOYQK 123  
 ID 033983 PRELIMINARY; PRT; 917 AA.  
 AC 033983;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE AMI.  
 GN Listeria monocytogenes.  
 OS Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 OX NCBI\_TaxID=1639;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=EGD;  
 RX MEDLINE=97426036; PubMed=9282740;  
 RA Braun L., Dramsi S., Dehoux P., Bienne H., Lindhal G., Cossart P.;  
 RT "Inlb : an invasion protein of Listeria monocytogenes with a novel  
 type of surface association.";  
 RL Mol. Microbiol. 25:285-294(1997).  
 DR EMBL; U082488; AAC45605.1; -  
 DR Interpro; IPR002502; Amidase\_2.  
 DR Pfam; PF01510; Amidase\_2; 1.  
 DR SMART; SM00644; Ami\_2; 1.  
 SQ SEQUENCE 917 AA; 102352 MW; 21EA2452B4F103B7 CRC64;

## RESULT 14

QY 64 YENDTSGNSQKCLQVOIEKEDDYTVSVETFRNASSPIKYYNTETKAVFOYQK 123  
 ID 033983 PRELIMINARY; PRT; 917 AA.  
 AC 033983;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE AMI.  
 GN Listeria monocytogenes.  
 OS Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 OX NCBI\_TaxID=1639;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=EGD;  
 RX MEDLINE=97426036; PubMed=9282740;  
 RA Braun L., Dramsi S., Dehoux P., Bienne H., Lindhal G., Cossart P.;  
 RT "Inlb : an invasion protein of Listeria monocytogenes with a novel  
 type of surface association.";  
 RL Mol. Microbiol. 25:285-294(1997).  
 DR EMBL; U082488; AAC45605.1; -  
 DR Interpro; IPR002502; Amidase\_2.  
 DR Pfam; PF01510; Amidase\_2; 1.  
 DR SMART; SM00644; Ami\_2; 1.  
 SQ SEQUENCE 917 AA; 102352 MW; 21EA2452B4F103B7 CRC64;

RL Arch. Pharm. Res. 24:316-322(2001).  
 DR EMBL: AY029222; AAK31641.1; -  
 KM Transference.  
 SO SEQUENCE 598 AA; 66501 MW; 4CE9BB26545F62A5 CRC64;

Query Match 7.7%; Score 87; DB 2; Length 598;  
 Best Local Similarity 24.0%; Pred. No. 28;  
 Matches 46; Conservative 25; Mismatches 55; Indels 66; Gaps 10;

QY 27 AGNPMAHE-----LGGKQDAMKSIQGVSVTVLAKTYEN-----66  
 DB 377 AGNN--WAHNSIAYAKDDSIILSSRHGCVN-IGRDQVKKWLAPSGWKNALASKLL 433  
 QY 67 -DTGSGWGFCKQVQVEIERKEDYTVT-----SVTFERN-----ASSPIK 106  
 DB 434 KPYDDGKGNALCKDCKNCENTDFDTYHTAMLSKGRITITDNGDGRGLEQPALPTMK 493  
 QY 107 Y-----YNTVE--TVKAVFOYG-----YKNIRNAIEYO-----VGGGLNTDTL 143  
 DB 494 YSHFVEYKIDKKGTQVQVWEYGERGYDFYSPITSVIEYQKDRDTMFGCGSINLFDVG 553  
 QY 144 IFPDGELCQVEY 155  
 DB 554 OPTIGKINETDY 565

## RESULT 9

QY 092N47 PRELIMINARY; PRT: 932 AA.

AC 092N47;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Putative outer membrane receptor protein.  
 GN R02383 OR SMC02721.  
 OS Rhizobium melioli (Sinorhizobium melioli).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Sinorhizobium.  
 OX NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1021.  
 RX MEDLINE=21396507; PubMed=11481430;  
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,  
 RA Bolstad P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,  
 RA Pohl T., Portetelle P., Puehler A., Purnelle B., Ramsperger U.,  
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;  
 RT "Analysis of the chromosome sequence of the legume symbiont  
 RT Sinorhizobium melioli strain 1021."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
 DR EMBL: AL591790: CAC46962.1; -;  
 DR InterPro: IPR000531; TonB\_boxC.  
 DR Pfam: PF00593; TonB\_dep\_Rec; 1.  
 KW complete proteome.  
 SO SEQUENCE 932 AA; 101872 MW; E4F6473BAF5AC484 CRC64;

Query Match 7.7%; Score 86.5; DB 16; Length 932;  
 Best Local Similarity 25.5%; Pred. No. 55;  
 Matches 41; Conservative 24; Mismatches 47; Indels 49; Gaps 11;

QY 44 DANKSIDGVSVTVLAKTYENDTG---SMGSO--FKCLQ--VOEIERKEDYTVTSVP 96  
 DB 221 DGRKIDRG-----GDTGDLTDMGKGLGFRADSTQNFERYQVYDETSNE 267  
 QY 97 TERNAS-----SPIRYN-----VETKAVFOYGYKKNRAIEYVGGGLNTDTLIF 145  
 DB 268 TYGLTLDDKDPFRFRYNASQKVIDAHNITQFSH-----SEYIPDINLTTVYR 320  
 QY 146 TDG-----ELCDVEYVNPADGCELMWKKSHYKHPD-VCT 180  
 DB 321 TDTTRAMYLKLVN--RNDSDTG--VWSLSNIIADPOTYST 356

## RESULT 10

QY 08XMD0 PRELIMINARY; PRT: 399 AA.

AC 08XMD0;  
 DT 01-MAR-2002 (Tremblrel. 20, Created)  
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE Hypothetical protein CPE0759.  
 GN CPE0759.  
 OS Clostridium perfringens.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 OX NCBI\_TaxID=1502;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=13 / Type A;  
 RX PubMed=11792842;  
 RA Shimizu T., Ohtani K., Hiraoka H., Ohshima K., Yamashita A.,  
 RA Shiba T., Ogasawara N., Hattori M., Kohara S., Hayashi H.;  
 RT "Complete genome sequence of Clostridium perfringens, an anaerobic  
 RT flesh-eater."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).  
 DR EMBL: AP003188; BAB80465.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SO SEQUENCE 399 AA; 46802 MW; 61019F0075828815 CRC64;

Query Match 7.6%; Score 86; DB 16; Length 399;  
 Best Local Similarity 23.8%; Pred. No. 20;  
 Matches 30; Conservative 21; Mismatches 37; Indels 38; Gaps 4;

QY 42 YQDAMS---IDG-----VSVTYVLAITYENDTSGWGFCKQVQVEIER 85  
 DB 224 YEEAKSLYEISGFILMCKRSYSENTEMLSFRIETNSIKASKKFAKMLSMIDM-- 281  
 QY 86 KEEDYVTSVFERNASPIKRYN-----TEVKAVFOYGYKKNRAIEYVGGGL 137  
 DB 282 -----YDNSKIYKINELGIIIRLEIAEIKVHOYCEYENIGPILEYDKSHGM 329  
 QY 138 NITDTL 143  
 DB 330 NLETL 335

## RESULT 11

QY 051717 PRELIMINARY; PRT: 584 AA.

AC 051717;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE Hypothetical 63.2 kDa protein precursor.  
 OS Pseudomonas fluorescens.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=294;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DSM 7139;  
 RX MEDLINE=93050280; PubMed=7961472;  
 RA Schlatter A., Jendrossek D.;  
 RT "Molecular characterization of the extracellular poly(3-  
 RT hydroxyoctanoic acid) [P(3HO)] depolymerase gene of Pseudomonas  
 RT fluorescens GR3 and of its gene product."  
 RL J. Bacteriol. 176:7065-7073(1994).  
 DR EMBL: U10470; AAA64539.1; -;  
 DR Hypothetical protein; Signal.  
 FT SIGNAL 1 26  
 SO SEQUENCE 584 AA; 63249 MW; 3AA16FAE898D2CBF CRC64;

Query Match 7.6%; Score 86; DB 2; Length 584;  
 Best Local Similarity 27.1%; Pred. No. 33;  
 Matches 29; Conservative 15; Mismatches 51; Indels 12; Gaps 2;

RP SEQUENCE FROM N.A.  
 RC STRAIN-Rhode Island; Tissue-Salivary gland;  
 RA Valenzuela J.G., Francischetti I.M., Pham V.M., Garfield M.,  
 RA Mather T.N., Ribeiro J.M.C.,  
 RT "Exploring the Salivome of the Tick Vector of Lyme Disease, Ixodes  
 RT scapularis."  
 RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF483718; AAM3640.1; -  
 SQ SEQUENCE 311 AA; 36581 MW; D7E529FEFE4CFA9A CRC64;

Query Match 9.2%; Score 103; DB 5; Length 311;  
 Best Local Similarity 19.6%; Pred. No. 0.45;  
 Matches 59; Conservative 35; Mismatches 79; Indels 128; Gaps 12;

OY 12 VSAALATQAEFTTSAKAGENPLMAHEELGKYODAMKSIDGVSVTYLAKTYENDPTGSW 71  
 DB 13 VSAALAEVEFOSWDRAPNPNPDLNKKDGAQDAPRTTKFTANHSYLYISS-----GW 66  
 OY 72 GSO-----FKLOVQEIERRKEDYVTSVTFERNASPIKYYN-----VTEYKAVF 118  
 DB 67 GTREHYEDVRCLOVHSSDLN-----YTLKSNYTSKMYNRTSRKMSNSTOYVQAK 117  
 OY 119 Q--YGYKNR-----NAIEQV-----GGGLNI-----TDFLI 144  
 DB 118 QKYYSIENIMHGGQFQREVTSPNGTCYMINFLCSEGGCRHHQECWQKRWTKYSEKYV 177  
 OY 145 FTDELCDVYFVNADOG---CELM-----YKSH----- 171  
 DB 178 LFTSPLCYVNSLDDDEYESCEFLSDWIKKNVTIPQVYTIKEKDESDLEKSEERE 237  
 OY 172 -----YKHVPDYCTEVENFCA-----KDRKTYDIFN 198  
 DB 238 SYERKESVQCEEPRTANKTFTYLDLFLKELPSSCRVAFNLNCGPRKRYIDKDCDIXIN 297  
 OY 199 E 199  
 DB 298 E 298

## RESULT 6

O95WZ5 PRELIMINARY; PRT; 221 AA.  
 AC O95WZ5;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
 DE Histamine binding protein.  
 GN HBP.  
 OS Ixodes scapularis (black-legged tick) (Deer tick).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Ixodidae; Ixodes.  
 OX NCBI\_Taxid=6945;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Das S., Banerjee G., Deponte K., Marcantonio N., Kantor F.S.,  
 RA Fikrig E.;  
 RT "Salp2SD, an Ixodes scapularis antioxidant, is one of 14  
 RT immunodominant antigens in engorged tick salivary glands."  
 RL J. Infect. Dis. 184:0-0(2001).  
 DR EMBL: AF209913; AAK97816.1; -  
 SQ SEQUENCE 221 AA; 25359 MW; 7B995AC9C0AFA252 CRC64;

Query Match 8.7%; Score 97.5; DB 5; Length 221;  
 Best Local Similarity 23.4%; Pred. No. 0.89;  
 Matches 49; Conservative 43; Mismatches 80; Indels 37; Gaps 14;

OY 10 TFVSAALATQAEFTTSAKAGENPLMAHEELGKY--ODAMKSIDGVSVTYVLA 60  
 DB 31 TYGSTGTTTTCGTGAGATVTTAPEDP-----SKYEONATRYVE--MNTQWVK 80  
 OY 61 KTYVENDTGSWGSQFKLOVEIERKEDYVTSV-FTERNASSPIKYYNTET--VKAV 117  
 DB 81 WRVYD-VTDFSGNVPVGCENFRVMEKR---TPTNVSFGYRYKRSKN-SMETIDETLLIKDI 134

OY 118 FOYGYKNIRNAIEYVGGLNITDPLI-FTDELCDVYFVNADOG---CELMYK--SHY 172  
 DB 135 GEHGFPPVNMFORPIG---TATDNLVYSNVCTVLRIPPTNGEHRCDLMMANLPLS 191  
 OY 173 KHVDPCTEVENFVCARDKRTYDIFNEEC 201  
 DB 192 QETPDDCLNKFPEVC-NTQIYRYVYVPC 219

## RESULT 7

O8IDR7 PRELIMINARY; PRT; 972 AA.  
 AC O8IDR7;  
 DT 01-MAR-2003 (TREMBLrel. 23, Created)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Hypothetical protein.  
 GN PF13\_0230.  
 OS Plasmodium falciparum (isolate 3D7).  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_Taxid=36329;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,  
 RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,  
 RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.,  
 RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AL844509; CAD52552.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 972 AA; 119242 MW; 70471A6C0327C32 CRC64;

Query Match 7.9%; Score 89; DB 5; Length 972;  
 Best Local Similarity 19.8%; Pred. No. 35;  
 Matches 43; Conservative 33; Mismatches 81; Indels 60; Gaps 9;

OY 4 QVLLTFVSAALATQAEFTTSAKAGENPLMAHEELGKYODAMKSIDGV-----SVYV 58  
 DB 289 KICLFIYISCSYISQYIKHMKKEN-----VEKKRKMKSIDNFIKINECKHI 340  
 OY 59 LAKTYENDTGSWGSQFKLOVQEI-----RKEEDYVTSVTFERNASSPIKYYNTETV 114  
 DB 341 LMDNLN-----KCINDYIRYIILRRKCSYLNINVLKTKLNSIKYLIINIK 390  
 OY 115 KAVQ--YGYKN-----IRATYQYVGGLNITDPLIFDGLCD-VFVFNADOGCELMVK 168  
 DB 391 KSFYICLYLNSYIFNKSITY-----CSDIFL-----LQKY 423  
 OY 169 KSHYKHVPDYCTEVENFVCARDKRTYDIFNECYVNG 205  
 DB 424 NLHTYIQLILLTFNLFKDKRKIYDYNKHLILYG 460

## RESULT 8

O939L2 PRELIMINARY; PRT; 598 AA.  
 AC O939L2;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Arylsulfate sulfotransferase.  
 GN ASTA.  
 OS Citrobacter freundli.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Citrobacter.  
 OX NCBI\_Taxid=546;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-M84-8242;  
 RX MEDLINE=21425731; PubMed=11534764;  
 RA Kang J.W., Jeong Y.J., Kwon A.R., Yun H.J., Kim D.H., Choi E.C.,  
 RT "Cloning, sequence analysis, and characterization of the asta gene  
 RT encoding an arylsulfate sulfotransferase from Citrobacter freundli.";

QY 181 FVFNVCARDKRTYDIFNEECYVNGEPM 209  
DB 181 FVFNVCARDKRTYDIFNEECYVNGEPM 209

## RESULT 2

Q8MV98 PRELIMINARY; PRT; 210 AA.

AC 08MV98; 01-OCT-2002 (Tremblrel. 22, Created)  
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
DE Putative 22.5 kDa secreted protein.  
OS Ixodes scapularis (Black-legged tick) (Deer tick).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Parasitiformes; Ixodidae; Ixodidae; Ixodes.  
OX NCBI\_Taxid=6945;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Rhode Island; TISSUE-Salivary gland;  
RA Valenzuela J.G., Francischetti I.M., Pham V.M., Garfield M.,  
Mather T.N., Ribeiro J.M.C.;  
RT "Exploring the Salivome of the Tick Vector of Lyme Disease, Ixodes  
scapularis."  
RU Submitted (Feb-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL:AF483742; AM93664.1; -  
SQ SEQUENCE 210 AA; 24689 MW; C5B75614649A9BED CRC64;

Query Match 11.2%; Score 126; DB 5; Length 210;

Best Local Similarity 24.3%; Pred. No. 0.0025;  
Matches 53; Conservative 41; Mismatches 90; Indels 34; Gaps 12;

QY 6 VLLTFVSAALATQAEFTSAKAGENPLMAHELLGKYODAMKSIDGVSVTYVLAQTYE 65  
DB 3 VMHFSFLCLLILVLD---AKGEIRI-DEDEKYMOYODIORALNPDRESMLYR-TYR 57  
QY 66 NIDGSGSOFKCL--OVQELERKEEDYTVTSVTFPNASSPIKYVVTETKAVAPQYK 123  
DB 58 REDD--GSEHICSAVSEHNPNGSDYEVQERLCT---KEQTRKTV-TLVATPYK 109  
QY 124 NRNALEYGVGGGLNTD-----TLIFTDELCDVFYV--PNADGCELMV-KKS 170  
DB 110 TENHATQRQNNMMARMSQKKDAQDGKYLIDYKCDILRLRNSGHDCLYLHSM 169  
QY 171 HKHNPDYCTFYFNVCARDKRTYDIFNEECYVNGEPM 208  
DB 170 LDDGVPRCESVYGIACGRDEPSY---KQRYV-YPM 201

## RESULT 3

Q8MV98 PRELIMINARY; PRT; 306 AA.

AC 08MV98; 01-OCT-2002 (Tremblrel. 22, Created)  
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
DE Putative secreted histamine binding protein.  
OS Ixodes scapularis (Black-legged tick) (Deer tick).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Parasitiformes; Ixodidae; Ixodidae; Ixodes.  
OX NCBI\_Taxid=6945;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Rhode Island; TISSUE-Salivary gland;  
RA Valenzuela J.G., Francischetti I.M., Pham V.M., Garfield M.,  
Mather T.N., Ribeiro J.M.C.;  
RT "Exploring the Salivome of the Tick Vector of Lyme Disease, Ixodes  
scapularis."  
RU Submitted (Feb-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL:AF483742; AM93664.1; -  
SQ SEQUENCE 306 AA; 34232 MW; B2DB09181CC4395D CRC64;

Query Match 10.4%; Score 117.5; DB 5; Length 306;

Best Local Similarity 20.0%; Pred. No. 0.023;  
Matches 46; Conservative 34; Mismatches 73; Indels 77; Gaps 7;

QY 1 MKQVVLITFVSAALATQAEFTSAKAGENPLMAHELLGKYODAMKSIDGVSVTYVLA 60  
DB 1 MGLMYAALFACIASASAMQAMTIRPENPL-LNPFSLGIPQSAMKALKKSKDDPFVLM 59  
QY 61 -----KTYENDTSGWSQFCKLOVOEIERKEEDYT 91  
DB 60 FRSRHNHPNITGVVYATSLHNETLKIVNFTFTYNET---DKNPTLEVOVALNQTQYK 116  
QY 92 VNSV-----FPRNASSPIKYVNTETKAVFOYGYKNI 125  
DB 117 LENVIRAGLKGPSDKPPLGNSMYEYDYSNTSSKPL-----TDMKAA----- 163  
QY 126 RNAIEYOYVG-----GLNITDPLIFTDGELCDVFYVNPADGCELMVKS 170  
DB 164 KDAVSGEASAPVEGVNLDIFYVHNOHCHILRSLPLKGGCDFLWRS 213

## RESULT 4

Q8I9T9 PRELIMINARY; PRT; 176 AA.

AC 08I9T9; 01-MAR-2003 (Tremblrel. 23, Created)  
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
DE Putative secreted histamine binding protein.  
OS Ornithodoros savignyi.  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Parasitiformes; Ixodidae; Argasidae; Ornithodoros.  
OX NCBI\_Taxid=69826;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Salivary gland;  
RX MEDLINE-21317941; PubMed-11425229;  
RA Mans B.J., Venter J.D., Vrey P.J., Louw A.I., Nelitz A.W.;  
RT "Identification of putative proteins involved in granule biogenesis of  
tick salivary glands."  
RU Electrophoresis 22:1739-1746(2001).  
DR EMBL:AF452891; AAN76831.1; -  
SQ SEQUENCE 176 AA; 19375 MW; 44B7595F3BCB6A25 CRC64;

Query Match 9.9%; Score 111.5; DB 5; Length 176;

Best Local Similarity 19.2%; Pred. No. 0.038;  
Matches 30; Conservative 33; Mismatches 64; Indels 29; Gaps 6;

QY 44 DAMKSIDGVSVTYVLAQTYENDTSGWSQFCKLOVOEIERKEEDYTVSVTFPNASS 103  
DB 23 DVM-NVLKGSDSKFLVAKRTTERGAN-----KCYMKRTSHDESHLTLEVLGYSKAGT 75  
QY 104 PIKY-----YNTETKAVFOYGYKNIR-----NAIEYGVGGGLNTDPLIFTDGELCDV 153  
DB 76 TDFVPSKYVTVYATSEGASTYNNMTYVRGSPASHGVKFE-----LVYSDGCGNI 125  
QY 154 FYVPA--DQCELMVKKSHYKHPDYCTFVNVFC 187  
DB 126 LQMKTSFPFGKCELMAPBEKARNVSSCGKFEKLC 161

## RESULT 5

Q8MV98 PRELIMINARY; PRT; 311 AA.

AC 08MV98; 01-OCT-2002 (Tremblrel. 22, Created)  
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
DE Putative secreted histamine binding protein.  
OS Ixodes scapularis (Black-legged tick) (Deer tick).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Parasitiformes; Ixodidae; Ixodidae; Ixodes.  
OX NCBI\_Taxid=6945;  
RN [1]

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2003, 17:13:11 ; Search time 95 seconds  
(without alignments)  
567.716 Million cell updates/sec

Title: DRET6  
Perfect score: 1125  
Sequence: 1 MKMQVLLTFVSALATQA.....DRKTYDFNECVNGEPWL 209

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1125	100.0	209	5	08WSK7
2	126	11.2	210	5	08MV98
3	117.5	10.4	306	5	08MV93
4	111.5	9.9	176	5	081979
5	103	9.2	311	5	08MV92
6	97.5	8.7	221	5	095W25
7	89	7.9	972	5	081DR7
8	87	7.7	598	2	093912
9	86.5	7.7	932	16	092M47
10	86	7.6	399	16	08XMD0
11	85.5	7.6	584	2	051717
12	85.5	7.6	917	2	033983
13	85.5	7.6	6761	2	081C77
14	85	7.6	437	10	09SQX6
15	85	7.6	476	17	08ZWE9
16	84.5	7.5	524	12	09EN29

17	84.5	7.5	917	2	050272	050272 listeria mo
18	84.5	7.5	917	16	08VA96	08VA96 listeria mo
19	84.5	7.5	1801	5	08WSJ2	08WSJ2 bombyx mori
20	84	7.5	218	5	095W78	095W78 ixodes scap
21	84	7.5	386	13	08AVY2	08AVY2 xenopus lae
22	83.5	7.4	482	5	002158	002158 caenorhabdi
23	83.5	7.4	1133	12	08JRX7	08JRX7 phthorimaea
24	83.5	7.4	2879	5	081C57	081C57 plasmodium
25	83	7.4	644	16	08EN80	08EN80 corynebacte
26	83	7.4	2285	9	064046	064046 bacterioph
27	83	7.4	2285	16	031976	031976 bacillus su
28	82	7.3	163	12	097D28	097D28 clostridium
29	82	7.3	360	16	098544	098544 paramedum
30	82	7.3	390	16	09PQ05	09PQ05 ureaplasma
31	82	7.3	561	5	09M4D7	09M4D7 caenorhabdi
32	82	7.3	564	5	095V41	095V41 caenorhabdi
33	82	7.3	637	2	0937B8	0937B8 pseudomonas
34	82	7.3	713	2	09JN39	09JN39 mycoplasma
35	82	7.3	876	5	09W034	09W034 drosophila
36	82	7.3	930	2	059290	059290 clostridium
37	82	7.3	1210	5	081AY1	081AY1 plasmodium
38	82	7.3	1604	2	09KR39	09KR39 rickettsia
39	81.5	7.2	370	2	09S003	09S003 borrelia bu
40	81.5	7.2	1092	16	08ZM12	08ZM12 salmonella
41	81	7.2	434	16	097N05	097N05 clostridium
42	80	7.1	262	2	087818	087818 roseobacter
43	80	7.1	3343	12	091HK6	091HK6 papaya ring
44	79.5	7.1	267	16	08RE04	08RE04 thermoaer
45	79.5	7.1	310	17	08TQ07	08TQ07 methanosarc

## ALIGNMENTS

RESULT 1  
ID 08WSK7 PRELIMINARY; PRT: 209 AA.  
AC 08WSK7;  
DT 01-MAR-2002 (TREMblrel. 20, Created)  
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)  
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
DE Serotonin and histamine binding protein.  
OS Dermacentor reticulatus.  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Parasitiformes; Ixodida; Ixodidae; Dermacentor.  
OX NCBI\_Taxid=57047;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Sangamadech S., Paesen G.C., Nuttall P.A.;  
RT "A high affinity serotonin- and histamine-binding lipocalin secreted  
RT by blood-feeding ticks.";  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF217101; AAL56644.1;  
DR InterPro; IPR002970; His\_binding.  
DR Pfam; PF02098; His\_binding; 1.  
SQ SEQUENCE 209 AA; 23889 MW; E5EB8BE710859591 CRC64;

Query Match 100.0%; Score 1125; DB 5; Length 209;  
Best Local Similarity 100.0%; Pred. No. 5; 6e-92;  
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKMQVLLTFVSALATQAETTSKAGENPMAHEELGKXODAMKSIDQGVSVTVYLA 60  
DB 1 MKMQVLLTFVSALATQAETTSKAGENPMAHEELGKXODAMKSIDQGVSVTVYLA 60  
QY KTYENDTGSMSQFCKLQVQETEKREDYTVSVFFRNASSPIKRYNTEVYKAVFY 120  
DB 61 KTYENDTGSMSQFCKLQVQETEKREDYTVSVFFRNASSPIKRYNTEVYKAVFY 120  
QY 121 GYKNIRNAIEYQVGGINITDTLFTDGLCDVFPYNADQGCGLMWKSKHYKHPDYCT 180  
DB 121 GYKNIRNAIEYQVGGINITDTLFTDGLCDVFPYNADQGCGLMWKSKHYKHPDYCT 180



Mon Aug 4 17:50:34 2003

dret6.rsp

Page 10

Search completed: August 4, 2003, 17:15:10  
Job time : 25 secs

---

OY 81 OEIERKEDYTVTSVFERNASSPIKYVNTETVKAVFOYGYKNINNAIEYOYGGGLNT 140  
 DB 210 -----KGLMALTAGRKOKHAKKVKYREFPASTSGVCWNPDLDELRLGTEYEGDVONIA 264  
 OY 141 DFLIFDGLCDVFFV-----PNADGCE 164  
 DB 265 TLIVDPDGKGLDSHMQKTAFFALLVGYLHATYKAKDGGTATLPSVDAMLADPNRIG-E 323  
 OY 165 LMKKSHRYKV 175  
 DB 324 LWMEMATYGVH 334

RESULT 14  
 DNK\_RICPR STANDARD: PRT: 627 AA.  
 AC Q9ZDX9;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Chaperone protein dnk (Heat shock protein 70) (Heat shock 70 kda protein) (HSP70).  
 GN DNK OR Rp185.  
 OS Rickettsia prowazekii.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;  
 CC Rickettsiaceae; Rickettsieae; Rickettsia.  
 OX NCBI\_Taxid-782;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Madrid E;  
 RX MEDLINE-99039499; PubMed-9823893;  
 RA Andersson S.G.E., Zomrodipour A., Andersson J.O.,  
 RA Sichteritz-Ponten T., Alsmark U.C.M., Podowski R.M., Neeslund A.K.,  
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;  
 RT "The genome sequence of Rickettsia prowazekii and the origin of  
 RT mitochondria";  
 RL Nature 396:133-140(1998).  
 CC -1- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).  
 CC -1- INDUCTION: By stress conditions e.g. heat shock (by similarity).  
 CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AJ235270; CAA14651.1; -  
 CC DR PIR: D71729; D71729.  
 CC DR HSSP: P04475; 1D64.  
 CC DR HAMAP: ME-00332; -; 1.  
 CC DR InterPro: IPR001023; Hsp70.  
 CC DR Pfam: PF00012; HSP70; 1.  
 CC DR PRINTS: PR00301; HEATSHOCK70.  
 CC DR PRODOM: P000089; HSP70; 1.  
 CC DR PROSITE: PS00297; HSP70\_1; 1.  
 CC DR PROSITE: PS00329; HSP70\_2; 1.  
 CC DR PROSITE: PS01036; HSP70\_3; 1.  
 CC KW Chaperone; ATP-binding; Heat shock; Phosphorylation;  
 KW Complete proteome.  
 FT MOD\_RES 197 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 SQ SEQUENCE 627 AA; 68383 MW; A638A0A5A5C8ACCT CRC64;

Query Match 7.2%; Score 81.5; DB 1; Length 627;  
 Best Local Similarity 28.2%; Pred. No. 12;  
 Matches 33; Conservative 15; Mismatches 42; Indels 27; Gaps 7;

OY 72 GSQFKCLOVQ-----IERKEDYTVTSVFERNASSPIKYVNTETVKAVFOYGY 121  
 DB 10 GTTNSCAVAMEGKEPKVIDNAEGERTPSIIAFANSERLVGOPAKRQAVT----- 59

OY 122 YKNINNAIEYOYGG--GLNITDPLIFDGLCDVFFVVPNADGCELMKKSHRYKV 176  
 DB 60 --NPRFTI-YAVKRLIGRNFETDPMWRKQGLVP-INIYKADNG-DAWEADNHRKSP 111

RESULT 15  
 HS9C\_DICDI STANDARD: PRT: 700 AA.  
 ID HS9C\_DICDI  
 AC P54651;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Heat shock cognate 90 kda protein.  
 GN HSPD OR HSC90.  
 OS Dictyostellium discoideum (Slime mold).  
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostellium.  
 OX NCBI\_Taxid-44689;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-AX3;  
 RA Boyes H., Ditttrich W., Mintert U., Lottspeich F., Gerisch G.,  
 RA Falx J.;  
 RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: MOLECULAR CHAPERONE. HAS ATPASE ACTIVITY  
 CC (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L43591; AAA69917.1; -  
 CC DR HSSP: P07900; 1YER.  
 CC DR DictyDB: DD00051; hspd.  
 CC DR InterPro: IPR003594; ATPbind\_ATPase.  
 CC DR InterPro: IPR001404; Hsp90.  
 CC DR Pfam: PF02518; HATPase\_C; 1.  
 CC DR Pfam: PF00183; HSP90; 1.  
 CC DR PRINTS: PR00775; HEATSHOCK90.  
 CC DR SMART: SM00387; HATPase\_C; 1.  
 CC DR PROSITE: PS00298; HSP90; 1.  
 CC KW Chaperone; ATP-binding; Heat shock.  
 KW SEQUENCE 700 AA; 79881 MW; BFB5ECF238089CD3 CRC64;

Query Match 7.2%; Score 81; DB 1; Length 700;  
 Best Local Similarity 19.5%; Pred. No. 15;  
 Matches 52; Conservative 32; Mismatches 105; Indels 78; Gaps 9;

OY 11 FVSALATQAEITTSKAG--ENPLWA-----HEELGK-----Y 42  
 DB 128 FYSAYLVADTYIVSHKNDDEQYVWESSAGGEFTALDHTPELGRTIVLHMKEDQDLY 187  
 OY 43 QDAWK-----SIQGVSVIYLAKTYYENDTSGWSGQFC-LOVOIERKEDY 91  
 DB 188 LDETIRKINLVKHSFEIYPIISLTIRKEKEVEDEETAKEGEESTDAKIEIEERKKK 247  
 OY 92 V---TSVTFERNASSPIKYVNTETVKAVFOYGYKNINNAIE-----YOVGGGLNTD 141  
 DB 248 VKVOEKEDVILKTKPLTRNPSDVTKEEYNSFYKSIISNDWEELAVAHFSVEGOLERKA 307  
 OY 142 TLIFDGLCDVFFVVPNADGCELMKKSHRYK-----VPDYCTFEVNFCAKD----- 190  
 DB 308 ILFVKKRNPFLPESKRRANNIKLVKRVFTIMDNCADLIIPYLNFRGIVSDPLNIS 367  
 OY 191 -----KTYDIFNE 199  
 DB 368 RETIQONKILTVIRKINLVKKEIELEFNE 394



```
QY 101 ASSPIKYVNTETKAVFQYGYKNIRNA-----IEYQVGGINTDTLIF 145
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 252 SKMALTKSKDEVSDESEYKEFKHIAHADDPLEVIAMKAEGEFERQ-----ALLF 302
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 146 TQGEICDVFYVFNADOGCELMWKKSHY-----KHVPDYCTFVFNFCARD 190
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 303 IPSHAFPDLFNSDAKIGMQLYVKRVFIMSDCDQMLPMYLRFFVKGVDAD 352
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
HTPG_RICPR ID HTPG_RICPR STANDARD; PRT; 621 AA.
AC 09ZCB9;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chaperone protein htpg (Heat shock protein htpg) (High temperature
  protein G).
GN HTPG OR RP840.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
  Rickettsiaceae; Rickettsiidae; Rickettsia.
OX NCBI_TaxId=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Madrid E;
RX MEDLINE-99039499; PubMed-9823893;
RA Andersson S.G.E., Zomrodipour A., Andersson J.O.,
  Ricketts A., Winkler H.H., Kurland C.G.;
  "The genome sequence of Rickettsia prowazekii and the origin of
  mitochondria."
RT Nature 396:133-140(1998).
RL
CC -1- FUNCTION: Molecular chaperone. Has ATPase activity (By
  similarity).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@sib-sib.ch).
CC
CC EMBL: AJ235273; CAA15264.1; -.
DR PIR: H71645; H71645.
DR HSSP: P02829; IAAH.
DR HAMAP: MF_00505; -.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR001404; Hsp90.
DR Pfam: PF02518; HATPase_C; 1.
DR Pfam: PF00183; HSP90; 2.
DR PRINTS: PRO0775; HEATSHOCK90.
DR SMART: SM00387; HATPase_C; 1.
DR PROSITE: PS00298; HSP90; 1.
KW Chaperone; ATP-binding; Heat shock; Complete proteome.
FT DOMAIN 1 328 A: SUBSTRATE-BINDING (BY SIMILARITY).
FT DOMAIN 329 544 B (BY SIMILARITY).
FT FT 545 621 C.
SQ SEQUENCE 621 AA; 70713 MW; 9F97B97801524007 CRC64;

Query Match 7.5%; Score 84.5; DB 1; Length 621;
Best Local Similarity 18.9%; Pred. No. 6.4;
Matches 39; Conservative 42; Mismatches 88; Indels 37; Gaps 8;

QY 11 FVSALALQAETTSKAGENPLMAHE-ELGKY--QDAMKSIDQGSVTVYIAK--TYIE 65
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 127 YSSFVADKVTYTSRKAGSKVHTWESDGLGIYVADSEQPTETGLVLYIKKSETTEL 186
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
QY 66 NDGWSGQFKCLQVQELERKEEDYTVTSFTFRNASSPIKYYNV-----TETKAVFQY 120
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 187 DH-----FRLKHIVKSYSDHIAVPIYCDGAGNNEIOLNSASALWTRPKSEITD 236
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 GYKNIRNAIEYQV-----GGLNTDTLIFTQGEICDVFYVFNADOGCELMWKK 169
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 237 QKFEFYKSLSYAVDDPWVTLHNKNEGCALIEFTMLTIPSRSKFDLPH-PDRKKRVLYLTKR 295
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 170 -----SHYKHVPDYCTFVFNFCARD 190
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 296 VFISDENIDLIPSLRFLRGVDSDD 321
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
Y692_METTH ID Y692_METTH STANDARD; PRT; 318 AA.
AC 026788;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MTH692.
GN MTH692.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
  Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxId=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Delta H;
RX MEDLINE-98037514; PubMed-9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
  Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
  Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
  Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
  Jiwani N., Caruso A., Bush D., Sater H., Patwell D., Prabhakar S.,
  McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
  Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
  "Complete genome sequence of Methanobacterium thermoautotrophicum
  deltaH: functional analysis and comparative genomics."
RT J. Bacteriol. 179:7135-7155(1997).
RL
CC -1- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@sib-sib.ch).
CC
CC EMBL: AE000848; AAB85197.1; -.
DR PIR: C69192; C69192.
DR InterPro: IPR001107; Band_7.
DR InterPro: IPR001972; Stomatin.
DR Pfam: PF01145; Band_7; 1.
DR PRINTS: PRO0721; STOMATIN.
DR SMART: SM00244; PHB; 1.
DR PROSITE: PS01270; BAND_7; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 2 22 POTENTIAL.
SQ SEQUENCE 318 AA; 35425 MW; 4A72C0AC8E99278D CRC64;

Query Match 7.3%; Score 82.5; DB 1; Length 318;
Best Local Similarity 23.3%; Pred. No. 4.3;
Matches 38; Conservative 23; Mismatches 77; Indels 25; Gaps 6;

QY 5 VLLITFVSALALQAETTSKAGENPLMAHEELGKYQDAMKSIDQGSVTVYIAKTY 64
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 10 VLVIAVAFKSLIKLIRYKGV-----ERGRKQ--RVESGLVLIIFIEIAIK 55
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 65 ENDGWSGQFKCLQVQELERKEEDYTVTSFTFRNASSPIK--YVNTETKAVFQYGYK 123
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 56 KVD---MREGVVDVDPQEVITKNDNTVVVDCVIFEVVDPPNAVNVVDFOALTKLAOT 112
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```



entites requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: AE001265; AAC65938.1; -

CC PIR: F71258; F71258.

DR HSSP: P02839; 1AH8.

DR TIGR: TP0964; -

DR HAMAP: MF\_00505; -; 1.

DR InterPro: IPR003594; Atptplnd\_Atpase.

DR InterPro: IPR001404; HSP90.

DR Pfam: PF02518; HATPase\_c.1.

DR Pfam: PF00183; HSP90\_2.

DR PRINTS: PR00775; HEATSHOCK90.

DR SMART: SM00387; HATPase\_c.1.

DR PROSITE: PS00298; HSP90; 1.

KM Chaperone; ATP-binding; Heat shock; Complete proteome.

FT DOMAIN 1 348 A: SUBSTRATE-BINDING (BY SIMILARITY).

FT DOMAIN 349 565 B (BY SIMILARITY).

FT DOMAIN 566 639 C.

SEQUENCE 639 AA; 72937 MW; 3E8FDBAC2282C31D CRC64;

Query Match 7.7%; Score 87; DB 1; Length 639;

Best local similarity 22.9%; Pred. No. 4;

Matches 50; Conservative 26; Mismatches 52; Indels 90; Gaps 10;

QY 5 VVLLTFVSALATQAEFTTSARAGENPLMAHEELLGKYQD-----AMKSIDQGV 54

DB 182 VVLLHLSQENSEFATR-----WRLEEVIAKKSVDHIAFPYLHLQEKYDKDGA 228

QY 55 VVYVIAKTYTENDTSMGQSFQCLQVQELERKEEDY-----IVTSVFPRNASSPIKYYN 109

DB 229 VVDTQKRDQVNDADALMKRPS-----ELKEEDYHREYQTLT-----RSTPPLLYVH 277

QY 110 VTEFKAAYFQYGYKRNIRNAIEYQVGGGLNITFTLFTDQ--ELCDVFYVP-----NA 159

DB 278 TK-----ASTGTQYVTLFTYPAKAPDPLDHA 303

QY 160 D--QGCELWVK-----SHYKHVPDYCFYFNVFCAND 190

DB 304 DYKPGVKLFVKRRFTTDEKELLPYLRFVRCVIDSED 341

RESULT 8

FAF\_DROME STANDARD: PRT; 2778 AA.

AC P55824; Q9Y9T6; Q9Y9T7; 01-NOV-1997 (Rel. 35, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DE 28-FEB-2003 (Rel. 41, Last annotation update)

DE Probable ubiquitin carboxyl-terminal hydrolase FAF (EC 3.1.2.15)

DE (Ubiquitin thiolesterase FAF) (Ubiquitin-specific processing protease FAF) (Deubiquitinating enzyme FAF) (Fat facets protein).

GN FAF OR BCDA:LD2582 OR CG1945.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Artthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3), AND TISSUE SPECIFICITY.

RC TISSUE=Eye, Imaginal disk;

RX MEDLINE=93202020; PubMed=1295747;

RA Fischer-Vize J.A., Rubin G.M., Lehmann R.;

RT "The fat facets gene is required for Drosophila eye and embryo development."

RL Development 116:985-1000(1992).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkley;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt J.G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Abmayant A., An H.-J., Andrews-Pfankoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,

RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brattler P.,

RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibbegan C.,

RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nussken D.R., Pacleb J.M.,

RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Port V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zheng G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster."

RL Science 287:2185-2195(2000).

[13]

RP REVISIONS, AND ALTERNATIVE SPLICING.

RC STRAIN=Berkley;

RX MEDLINE=22426069; PubMed=12537572;

RA Mitra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,

RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochuk S.E.,

RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,

RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,

RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,

RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,

RA Lewis S.E.;

RT "Annotation of the Drosophila melanogaster euchromatic genome: a

RT systematic review."

RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).

[14]

RP SEQUENCE OF 1089-2778 FROM N.A. (ISOFORM 1).

RC STRAIN=Berkley;

RX MEDLINE=20196012; PubMed=10731138;

RA Rubin G.M., Hong L., Brokstein P., Evans-Holm M., Erise E.,

RA Stapleton M., Harvey D.A.;

RT "A Drosophila complementary DNA resource."

RL Science 287:2222-2224(2000).

-I- FUNCTION: REQUIRED FOR EYE AND EMBRYO DEVELOPMENT, AND PLAYS A

ROLE IN COMPOUND EYE ASSEMBLY AND OOGENESIS RESPECTIVELY. IN THE

LARVAL EYE DISKS, CELLS OUTSIDE THE ASSEMBLING FACETS REQUIRE THIS

PROTEIN FOR SHORT-RANGE CELL INTERACTIONS THAT PREVENT THE MYSTERY

CELLS FROM BECOMING PHOTORECEPTORS. IT IS ALSO REQUIRED FOR

NUCLEAR MIGRATION AND CELLULARIZATION IN EARLY EMBRYOGENESIS AND

COULD PLAY A ROLE IN POLE CELL DETERMINATION, DEVELOPMENT OR

FUNCTION.

-I- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =

ubiquitin + a thiol.

-I- ALTERNATIVE PRODUCTS:

Event-Alternative splicing: Named isoforms=3:

Comment-Experimental confirmation may be lacking for some

isoforms;

Name=1;

Isoid=P55824-1; Sequence=Displayed;

```

Db      183 DT--FLOHFR---LKHIVKSYSDHIANVPIYFFDEAGNNEIDNBSASALMTPRKSEITBEO 237
Oy      118 FOYGYKKNIRNAIE-----YOVGGGLNTDTLFTDGEICDVFYVVPNADGCELMWKK- 169
Db      238 YKEEYKSIYSYADDPWITMHNKNGCAIEFTMLTFLPSSKTFDLFH-PDKRRRVKLYIKRV 296
Oy      170 ----SHYKNVPDYCTFVFNVCARD 190
Db      297 FISPENIDLPISYLFRLRGVDSIED 321

RESULT 6
HTPG_BORBU
ID      HTPG_BORBU      STANDARD:      PRT:      616 AA.
AC      P42535.
DT      01-NOV-1995 (Rel. 32, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Chaperone protein htpg (Heat shock protein htpg) (High temperature
DE      protein G).
DE      HTPG OR B0560.
OC      Borrelia burgdorferi (Lyme disease spirochete).
OC      Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX      NCBI_TaxID=139;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=297;
RA      Porcella S.F., Radolf J.D., Norgard M.V.;
RL      Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 35210 / B31;
RA      Fraser C.M., Castens S., Huang W.M.;
RA      MEDLINE=98065943; PubMed=9403685;
RA      Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwin M.,
RA      Dougherty J., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA      Peterson S., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA      van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman C.,
RA      Uterback T., Matthey L., McDonald L., Artlach P., Bowman C.,
RA      Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA      Smith H.O., Venter J.C.;
RP      "Genomic sequence of a Lyme disease spirochaete, Borrelia
RP      burgdorferi."
RL      Nature 390:580-586(1997).
RN      [3]
RP      SEQUENCE OF 497-574 FROM N.A.
RC      STRAIN=212;
RA      MEDLINE=95111614; PubMed=7812434;
RA      Ojalim C., Davidson B.E., Saint-Gilons I., Old I.G.;
RT      "Conservation of gene arrangement and an unusual organization of rRNA
RT      genes in the linear chromosomes of the Lyme disease spirochaetes
RT      Borrelia burgdorferi, B. garinii and B. afzelii."
RL      Microbiology 140:2931-2940(1994).
CC      -1- FUNCTION: Molecular chaperone. Has ATPase activity (By
CC      similarity).
CC      -1- SUBUNIT: Homodimer (By similarity).
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC      -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: U51878; AAA97469.1; -
DR      EMBL: AE001157; AAC66919.1; ALT_INIT.
DR      EMBL: L32145; AAC41403.1; -
DR      HSSP: P07900; IYER.
DR      TIGR: B0560; -.

```

Query Match	8.5%	Score 96	DB 1	Length 616
Best Local Similarity	21.5%	Pred. NO. 0.61		
Matches 46	Conservative 34	Mismatches 92	Indels 42	Gaps 8
OY	11	FVSAALATQ-AETTSARAGEMP--LMAHEELGKYQDAWKSIDGVSTVYLAKTTEEND 67		
DB	121	FYSAFIYSEKEVEYNSKNALESDAIWMSDGKTGYEIAKAKEESGTEKLNLNKEGLE-- 178		
OY	68	TGSMGSOFCLOV-----QIERKEDYVTSVTFPRNA 101		
DB	179	--YANKWKLOEIIKKTSNMHINYPYIIKSEPIPKDKGKOEIEEKEELNETTALMTKN- 234		
OY	102	SSPIKYYNVETFAVPOGYGKKRNRALEYGVGGGLNTDPLFTFDGLCVFVPPNADQ 161		
DB	235	KSEIKAEYENEFYKNT--TFDYENPLMHITHTAEGNLETTNLFYVPSKAPDYLIY-PNTKP 292		
OY	162	GCELMVVK-----SHYKHVPDYCTFENVFVCAKD 190		
DB	293	GVKLFINRIFTDSEGSLLPWYLFRIKGIIDCOD 326		
RESULT 7				
HTPG-TREPA				
ID	HTPG-TREPA	STANDARD:	PRT:	639 AA.
AC	O83949:			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Chaperone protein htpg (Heat shock protein htpg) (High temperature protein G).			
GN	HTPG OR TP0984.			
OS	Treponema pallidum.			
OC	Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.			
OX	NCBI_TaxID=160;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Nichols;			
RX	MEDLINE=98332770; PubMed=9665876;			
RA	Frisser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,			
RA	Dodson R., Gwyn M., Hickey E.K., Clayton R., Ketchum K.A.,			
RA	Sodergren E., Hardham J.M., McLeod M.P., Seitzberg S., Peterson J.,			
RA	Khalak H., Richardson D., Howell J.K., Chidambaram M., Utechtack T.,			
RA	McDonald L., Artlach P., Bowman C., Cotton M.D., Fujii C., Garland S.,			
RA	Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,			
RA	Venter J.C.;			
RT	"Complete genome sequence of Treponema pallidum, the syphilis			
RT	spirochete."			
RL	Science 281:375-388(1998).			
CC	-1- FUNCTION: Molecular chaperone. Has ATPase activity (By			
CC	similarity).			
CC	-1- SUBUNIT: Homodimer (By similarity).			
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).			
CC	-1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation --			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			

```

FT CHAIN 19 190 FEMALE-SPECIFIC HISTAMINE-BINDING PROTEIN
FT DISULFID 66 187 1.
FT DISULFID 137 166 BY SIMILARITY.
SQ SEQUENCE 190 AA: 21370 MW: 855BE151A9003B1 CRC64;

Query Match
Best Local Similarity 22.4%; Score 252; DB 1; Length 190;
Matches 68; Conservative 28; Mismatches 89; Indels 18; Gaps 8;

QY 3 MQVALLTFVSAALTAQAEETSAKAGENPLMAHELLGKYDQAMKSIDQGVSYTVYLAKT 62
DB 1 MKLLSLAFLVALSQVKA-----DKPVMADEKANGEHODAMKHLQKLYEENYDLTKA 52
QY 63 TYENDTGSWGSQFKLQVQEIERRKEDYTVSVFTFRNASSPIKYYNTEYKAVFOYGY 122
DB 53 TYKNRP-VWGNDFTCVGAQAQNLNDEKNVEMFEMNADIV-YQHFEKATPPKMGY 110
QY 123 KNIRNAIEYQVGGGINTDTLFTDGLCDVFEYVNPAD--QGCELMVKKSHYKHVPDYC 179
DB 111 -NKENAIYQFEDGGVLLIDVLAFLSD-DNCYVYIALGPPGSGAGYELMA--TDYTDVPASC 166
QY 180 TFEVNFCAKDRKRTDIFNECV 202
DB 167 LEKENEYAA-GIIPRDVYTSDDL 188

RESULT 4
NGGC_MESVI STANDARD: PRT; 174 AA.
AC 09MURI;
DF 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NAD(P)H-quinone oxidoreductase chain J, chloroplast (EC 1.6.5.-)
DE (NAD(P)H dehydrogenase, chain J) (NADH-plastoquinone oxidoreductase
  subunit J).
GN NDHJ.
OS Mesostigma viride.
OC Chloroplast.
CC Eukaryota; Viridiplantae; Streptophyta; Mesostigmatophyceae;
  Mesostigmatales; Mesostigmataceae; Mesostigma.
OX NCBI_TaxID=41882;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-NIES-296;
RX MEDLINE=20150907; PubMed=10688199;
RA Lemieux C., Otis C., Turmel M.;
RT "Ancestral chloroplast genome in Mesostigma viride reveals an early
  branch of green plant evolution.";
RL Nature 403:649-652(2000).
CC -1- CATALYTIC ACTIVITY: NAD(P)H + plastoquinone = NAD(P)(+) +
  plastoquinol.
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I 30 kDa SUBUNIT FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AF161114; AAF43839.1; -
DR InterPro: IPR001268; Complex1_30K.
DR Pfam: PF00329; complex_30kd; 1.
DR PROSITE: PS001581; Complex1_30K; 1.
DR PROSITE: PS00542; COMPLEX1_30K; 1.
KW Oxidoreductase; NAD; NADP; Quinone; Plastoquinone; Chloroplast.
SQ SEQUENCE 174 AA: 20329 MW: 122ABFA9188D47B7 CRC64;

Query Match
Best Local Similarity 8.7%; Score 97.5; DB 1; Length 174;
Matches 27.0%; Pred. No. 0.097;

```

```

Matches 31; Conservative 20; Mismatches 37; Indels 27; Gaps 5;

QY 105 IKYVNTETVAVAPQYKKNRMIEYVGGGINTDTLFTDDELCDVFEV----PAD 160
DB 47 VKAOMLTVSAIALTAAGFNRLSQCAVDVSPG-----GDLASVHLTVDDNAD 95
QY 161 QGCELMVKKSHYKHVPDYCTFEVFNFCA-----KDRKTYDIFNECVNGEPML 209
DB 96 QPQEVCIK-----VFPFRKRPILPSVFWTKRADQERSYDMFG--IYBSHPHL 144

RESULT 5
HTPG_RICCN STANDARD: PRT; 621 AA.
AC P58478;
ID HTPG_RICCN
DF 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chaperone protein htpg (Heat shock protein htpg) (High temperature
  protein G).
GN HTPG OR R01302.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
  Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=781;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-Malish 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Remesto-Audiffren P., Fournier P.-E., Barbe V.,
  Samsom D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
  Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL Science 293:2093-2098(2001).
CC -1- FUNCTION: Molecular chaperone. Has ATPase activity (By
  similarity).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AE008676; AAL03840.1; -
DR PIR: F97862; F97862.
DR HAMAP: MF_00505; -; 1.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR001404; Hsp90.
DR Pfam: PF02518; HATPase_c; 1.
DR Pfam: PF00183; HSP90; 2.
DR PRINTS: PR00775; HEATSHOCK90.
DR SMART: SM00387; HATPase_c; 1.
DR PROSITE: PS00298; HSP90; 1.
KW Chaperone; ATP-binding; Heat shock; Complete proteome.
FT DOMAIN 1 328
FT DOMAIN 329 544 A; SUBSTRATE-BINDING (BY SIMILARITY).
FT DOMAIN 545 621 B (BY SIMILARITY).
SQ SEQUENCE 621 AA: 70804 MW: PF6E5B35DDBF017 CRC64;

Query Match
Best Local Similarity 8.6%; Score 96.5; DB 1; Length 621;
Matches 43; Conservative 46; Mismatches 81; Indels 35; Gaps 10;

QY 11 FVSAALTAQAEETSAKAGENP--LMAHELLGKY--QDAMKSIDQGVSYTVYLAKTYYEN 66
DB 127 YSSFWVADKVTIKVTIRKAGEDVHTM-ESDGETVSDSDKEFTRGTEIVLHKV---EE 182
QY 67 DTGSMGSGFKLQVQEIERRKEDYTVSVFTF-----RNASSPIKYYNTEYKAV 117

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2003, 17:13:07 ; Search time 24 seconds

(without alignments)  
409.524 Million cell updates/sec

Title: DRET6  
Perfect score: 1125  
Sequence: 1 MKMQVLLTTFVSALATQA.....DRKTYDIFNECVNGEPM 209

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	290.5	25.8	190	1	HBP2_RHIAP
2	265	23.6	200	1	HBPM_RHIAP
3	232	22.4	190	1	HBP1_RHIAP
4	97.5	8.7	174	1	NUCC_MESVI
5	96.5	8.6	621	1	HTPG_RICCN
6	96	8.5	616	1	HTPG_BORBU
7	87	7.7	639	1	HTPG_TREPA
8	86	7.6	2778	1	FAR_DROME
9	85	7.6	656	1	HTPG_MYCLE
10	84.5	7.5	621	1	Y692_METHH
11	82.5	7.3	318	1	Y692_METHH
12	82	7.3	381	1	RIR5_CAEEL
13	82	7.3	637	1	TRG5_ECOLI
14	81.5	7.2	627	1	DNAX_RICPR
15	81	7.2	700	1	HS9C_DICDI
16	80.5	7.2	621	1	HTPG_HELPJ
17	79.5	7.1	621	1	HTPG_HELFP
18	79.5	7.1	629	1	HTPG_CHLFE
19	78.5	7.0	864	1	Y63M_YEAST
20	78	6.9	215	1	NO22_SOYBN
21	77	6.8	523	1	HEX_ADEMI
22	77	6.8	909	1	HEX_ADEMI
23	76.5	6.8	393	1	IDH_STRMU
24	76.5	6.8	724	1	IDH_STRMU
25	76	6.8	260	1	DDX4_HUMAN
26	76	6.8	1031	1	YD9G_SCHPO
27	76	6.8	1068	1	YD9G_SCHPO
28	76	6.8	1068	1	YD9G_SCHPO
29	75.5	6.7	1489	1	YD9G_SCHPO
30	75	6.7	1489	1	YD9G_SCHPO
31	75	6.7	424	1	Y878_METUA
32	75	6.7	424	1	Y878_METUA
33	75	6.7	626	1	HTPG_BUCBP
			647	1	HTPG_MYCTU

34	75	6.7	3951	1	VGFL_IBVB
35	74.5	6.6	400	1	GTR3_RABIT
36	74.5	6.6	492	1	VGCE_ECOLI
37	74.5	6.6	726	1	NFT1_YEAST
38	74.5	6.6	890	1	LPN1_HUMAN
39	74.5	6.6	2768	1	THYG_HUMAN
40	74	6.6	406	1	IDH_SPHYA
41	74	6.6	854	1	TRIC_SULTO
42	74	6.6	912	1	TBPL_HAEIN
43	73.5	6.5	355	1	UBPC_HUMAN
44	73.5	6.5	553	1	GRP2_HUMAN
45	73.5	6.5	2896	1	HCYG_OCTDO

## ALIGNMENTS

RESULT 1	ID	HBP2_RHIAP	STANDARD:	PRT:	190 AA.
AC	077421:				
DT	30-MAY-2000 (Rel. 39, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Female-specific histamine-binding protein 2 precursor (FS-HBP2).				
OS	Rhipicephalus appendiculatus (Brown ear tick).				
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;				
OC	Parasitiformes; Ixodidae; Ixodidae; Rhipicephalus.				
OX	NCBI_TaxID=34631;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (1.25 ANGSTROMS).				
RC	TISSUE=Salivary gland;				
RX	MEDLINE=99288454; PubMed=10360182;				
RA	Paesen G.C., Adams P.L., Harlos K., Nuttall P.A., Stuart D.I.;				
RT	"Tick histamine-binding proteins: Isolation, cloning, and three-				
RL	dimensional structure.";				
CC	Mol. Cell 3:661-671(1999).				
CC	-1- FUNCTION: BINDS HISTAMINE WITH A HIGH-AFFINITY. THE ABILITY TO				
CC	OUTCOMPETE HISTAMINE RECEPTORS INDICATES THAT ITS FUNCTION IS TO				
CC	SUPPRESS INFLAMMATION DURING BLOOD FEEDING.				
CC	-1- SUBCELLULAR LOCATION: Secreted.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	-----				
CC	EMBL: U96081; AAC63107.1; -				
DR	PDB: 1QFT; 19-APR-00.				
DR	PDB: 1QFV; 19-APR-00.				
DR	InterPro: IPR002970; His_binding.				
DR	Pfam: PF02098; His_binding; 1.				
DR	ProDom: PD152455; His_binding; 1.				
KW	Signal; 3D-structure.				
FT	SIGNAL	1	19		
FT	CHAIN	20	190		
FT	DISULFID	67	188		
FT	DISULFID	138	167		
FT	TURN	23	24		
FT	HELIX	27	30		
FT	HELIX	31	33		
FT	HELIX	36	41		
FT	STRAND	42	45		
FT	STRAND	48	53		
FT	STRAND	57	58		
FT	TURN	59	61		
FT	STRAND	62	63		
FT	TURN	64	64		
FT	STRAND	66	76		

-----  
FEMALE-SPECIFIC HISTAMINE-BINDING PROTEIN  
2.

Mon Aug 4 17:50:32 2003

dret6.rpr

Page 6

```

OY      8I QELERREEDYTVSVTEPRNASSPLKYYNVTETVAAPQYGYKKNIRNAIEYQVGGGLNIT 140
        | : : : : : : : : | : : : : : : : : | : : : : : : : : |
Db      210 ----GGEIMALLTAGRORHAKKNVLRPEASTSGGVCWNPDLRIETGEYEVGDYONLA 264
        | : : : : : : : : | : : : : : : : : | : : : : : : : : |
OY      141 DTLIFFDGELCDVEFY-----PNAODGCE 164
        :: || : : : : : : : : | : : : : : : : : |
Db      265 TLIVDPDGKGLDSHWOKTFAILLVGYIHLALYKAKADGCTATLPSVDAMLADPNRDIG-E 323
        | : : : : : : : : | : : : : : : : : | : : : : : : : : |
OY      165 LAYKSKSHYKHV 175
        | : : : : : : : : | : : : : : : : : |
Db      324 LMHEMATYGHV 334

```

Search completed: August 4, 2003, 17:15:58  
Job time : 42 secs

```

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1381 <WILL>
A:Cross-references: EMBL:Z35637; PTDN:CAAB4688.1; GSPDB:GN00021; CESP:C03C10.3
A:Experimental source: clone C03C10
C:Genetics:
A:Gene: CESP:C03C10.3
A:Map position: 3
A:Introns: 18/3; 137/3; 256/2
C:Superfamily: herpesvirus ribonucleoside-diphosphate reductase small chain

Query Match
Best Local Similarity 7.3%; Score 82; DE 2; Length 381;
Matches 29; Conservative 20; Mismatches 56; Indels 28; Gaps 3;

OY 9 LTFVSALATQAEFTTSKAGENPLMAHELLG-----KYQDAKSIDGVSSTY 57
Db 35 LEIDQGTAAAEETNNSEVNEELDADEPMLODDNRFVLPFLKHHIDMFKKAVASFW 94

OY 58 VLAATTYENDGSMGSOFCLOVOEIERKEEDYTVTSVTFPRNASSPI-----KYNNV 110
Db 95 TVEEVDLGKDNMDW-----EKNMGDDQYIFISRLTAFPAASDGIWNEICERFSNE 144

OY 111 TETVKAAYQYGYK 123
Db 145 VQVSEARFEGFQ 157

RESULT 13
DB82922
conserved hypothetical UU198 [Imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: DB82922
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Helner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to Genbank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
A:Reference number: A82870
A:Accession: DB82922
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-390 <GLA>
A:Cross-references: GB:A8002119; GB:AF222894; NID:q6899153; PTDN:AAF30605.1; GSPDB:GN000
A:Experimental source: serovar 3; biovar 1
C:Genetics:
A:Gene: UU198
A:Genetic code: SGC3

Query Match
Best Local Similarity 7.3%; Score 82; DB 2; Length 390;
Matches 29; Conservative 28; Mismatches 51; Indels 22; Gaps 5;

OY 8 LTFVSALATQAEFTTSKAGENPLMAHELLGKYQDAKSIDGVSSTY-VLAATTYEN 66
Db 167 ILAVSRGGAASGTAEAFLLSNPLSKKEKKKHIDYKYIENG---TYGKLIRSTEES 223

OY 67 DTGSMGSOFCLOVOEIERKEEDYTVTSVTFPRNASSPIKYYNTEYKAVFYGY--K 123
Db 224 NTQAM-----TNISKNDVDSITYL-----STGFILKNYDOIIAKGYKTIAYLQNK 268

OY 124 NIRNAIEYOV 133
Db 269 DKNTLPPEI 278

RESULT 14
T08530
frag protein - Enterobacter aerogenes plasmid R751
C:Species: Enterobacter aerogenes
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: T08530
R:Thorsted, P.B.; Shah, D.S.; Macartney, D.; Kostelidou, K.; Thomas, C.M.
plasmid 36, 95-111, 1996

```

```

A>Title: Conservation of the genetic switch between replication and transfer genes of
A:Reference number: Z16434; MUID:971118926; PMID:8954881
A:Accession: J08530
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-637 <THO>
A:Cross-references: EMBL:X54458; NID:g1572520; PIDN:ACC64474.1; PID:g1572579
C:Genetics:
A:Gene: trag
A:Genome: plasmid R751

Query Match          7.3%; Score 82; DB 2; Length 637;
Best Local Similarity 18.3%; Pred. No. 25;
Matches 46; Conservative 30; Mismatches 89; Indels 86; Gaps 7;

OY 5 VLLITFVSALATQAETTSKAGE---NPLMAHEE----- 37
   : : : : : : : : : : : : : : : : : : : :
DB 90 LVSTVGLGVAIVAAKVYVNSSSKANEDYLHGASRAAEKKDIOAAGLLPRERVLEIVTGKAA 149
   : : : : : : : : : : : : : : : : : : : :
OY 38 --LLGKYDADAKSID-----OGSVTYVLAATTYENTGT-----SMGSGFCKLQY 80-
   : : : : : : : : : : : : : : : : : : : :
DB 150 PTAIVGVYGGQODKDNFETFLRHSGPEHVLTYAATPSRGKGVGLVPLPLSGASSVITDL 209
   : : : : : : : : : : : : : : : : : : : :
OY 81 QEIRKEEDYVTSVFETFRNASSPIKYYNVEYKAFQYGYKNIRNAIEYOGGGGLNIT 140
   : : : : : : : : : : : : : : : : : : : :
DB 210 -----KGLMALTLTGWRGKHAHKNKVLNFEPASTSGGVCMNPLDEIRLGTETEPYGDVGNLA 264
   : : : : : : : : : : : : : : : : : : : :
OY 141 DTLITFDGELCDVRYV-----LPLADGCGE 164
   : : : : : : : : : : : : : : : : : : : :
DB 265 TLIVDPDGKGDISHMOKTAPFALLGVILHALYKAKDDGTATLPSVDAMLADPNRIDG-E 323
   : : : : : : : : : : : : : : : : : : : :
OY 165 LMVKKSHYKHV 175
   : : : : : : : : : : : : : : : : : : : :
DB 324 LMMEKATYGHV 334

RESULT 15
S22992
trag protein - Escherichia coli plasmid R751
C:Species: Escherichia coli
C:Date: 31-Dec-1993 #sequence_revision 02-Aug-1994 #text_change 08-Oct-1999
C:Accession: S22992; S42274
R:Ziegelin, G.; Pansgrau, W.; Strack, B.; Balzer, D.; Kroege, M.; Kruft, V.; Lanka,
DNA Seq. 1, 303-327, 1991
A>Title: Nucleotide sequence and organization of genes flanking the transfer origin o
A:Reference number: S22992; MUID:92190548; PMID:1665997
A:Accession: S22992
A:Molecule type: DNA
A:Residues: 1-637 <ZIE>
A:Cross-references: EMBL:X54458
R:Lanka, B.
submitted to the EMBL Data Library, August 1990
A:Reference number: S42274
A:Accession: S42274
A:Molecule type: DNA
A:Residues: 1-628; 'D', 630-637 <LAN>
A:Cross-references: EMBL:X54458; NID:g42656; PIDN:CAA8327.1; PID:g42657
C:Genetics:
A:Gene: trag
A:Genome: plasmid

Query Match          7.3%; Score 82; DB 2; Length 637;
Best Local Similarity 18.3%; Pred. No. 25;
Matches 46; Conservative 30; Mismatches 89; Indels 86; Gaps 7;

OY 5 VLLITFVSALATQAETTSKAGE---NPLMAHEE----- 37
   : : : : : : : : : : : : : : : : : : : :
DB 90 LVSTVGLGVAIVAAKVYVNSSSKANEDYLHGASRAAEKKDIOAAGLLPRERVLEIVTGKAA 149
   : : : : : : : : : : : : : : : : : : : :
OY 38 --LLGKYDADAKSID-----OGSVTYVLAATTYENTGT-----SMGSGFCKLQY 80
   : : : : : : : : : : : : : : : : : : : :
DB 150 PTAIVGVYGGQODKDNFETFLRHSGPEHVLTYAATPSRGKGVGLVPLPLSGASSVITDL 209
   : : : : : : : : : : : : : : : : : : : :
OY 81 QEIRKEEDYVTSVFETFRNASSPIKYYNVEYKAFQYGYKNIRNAIEYOGGGGLNIT 140
   : : : : : : : : : : : : : : : : : : : :
DB 210 -----KGLMALTLTGWRGKHAHKNKVLNFEPASTSGGVCMNPLDEIRLGTETEPYGDVGNLA 264
   : : : : : : : : : : : : : : : : : : : :
OY 141 DTLITFDGELCDVRYV-----LPLADGCGE 164
   : : : : : : : : : : : : : : : : : : : :
DB 265 TLIVDPDGKGDISHMOKTAPFALLGVILHALYKAKDDGTATLPSVDAMLADPNRIDG-E 323
   : : : : : : : : : : : : : : : : : : : :
OY 165 LMVKKSHYKHV 175
   : : : : : : : : : : : : : : : : : : : :
DB 324 LMMEKATYGHV 334

150 PTAIVGVYGGQODKDNFETFLRHSGPEHVLTYAATPSRGKGVGLVPLPLSGASSVITDL 209

```

A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yomi

Query Match 7.4%; Score 83; DB 2; Length 2285;  
Best Local Similarity 23.3%; Pred. No. 1.1e+02;  
Matches 51; Conservative 27; Mismatches 79; Indels 62; Gaps 10;

OY 6 VLLTTFVSALATQAETTSKAGENPLMAHEELLGKODAKSID--QGVSVYVLAKT 63  
DB 743 LVTSTLVGAFALG-----MALESIISSFAEKAKKDDFEQSQTNVERIT 790  
OY 64 YENDTGSWSGQFCLQVOEIERKEEDYTVTSFTFRNASSPIKYNTETVKAFFQ--Y 120  
DB 791 NKDSTDKLIQYK-----ELQVKESRSLTS-----DEQGYLQVTLQQLAQTFPALVK 838  
OY 121 GY-----KNIRNAIEYGVGGGLNTDPLFTDGLCDVFPVPNADGCELMVK 169  
DB 839 GYDSQGNALIKTKNELEKALE-----NTRKEYLALKKQETRDŠAKKTFEDASKE--IKK 889  
OY 170 S-----HYKHVPDYCTFVFVNFPCAKDRKTYDIFNEECVY 203  
DB 890 SKDELKQYKQIADYND-----KGRPKMDLIADDDY 920

## RESULT 9

C69192  
stomatin-like protein - Methanobacterium thermoautotrophicum (strain Delta H)  
C:Species: Methanobacterium thermoautotrophicum  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 24-Sep-1999

C:Accession: C69192  
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
Qiu, D.; Spedaleora, R.; Viciore, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jilani, N.;  
Kl, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
A:Reference number: A69000; MWID:98037314; PMID:93711463  
A:Accession: C69192  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-318 <MTH>  
A:Cross-references: GB:AE000848; GB:AE000666; NID:g2621761; PIDN:AA85197.1; PID:g262177  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH692  
C:Superfamily: erythrocyte band 7 integral membrane protein

Query Match 7.3%; Score 82.5; DB 2; Length 318;  
Best Local Similarity 23.3%; Pred. No. 9.2;  
Matches 38; Conservative 23; Mismatches 77; Indels 25; Gaps 6;

OY 5 VLLTTFVSALATQAETTSKAGENPLMAHEELLGKODAKSIDQGVSVTVLAKT 64  
DB 10 VLLVLAFLKSLKILRPYKGV-----ERLGRYQ---RTVEGLVITIIPTFAIK 55  
OY 65 ENDTGSWSGQFCLQVOEIERKEEDYTVTSFTFRNASSPIK--YVNTETVKAFFQYK 123  
DB 56 KVP---MREQVVPVPPQEVITKDNVTVVVDCVIFEEVDPFNNAVYNVDEYQAITKLAQT 112  
OY 124 NINNAIEYGVGGGLNTDPLFTDGLCDVFPVPNADGCELMVK 166  
DB 113 NLNRII-----GDELDQTL--TSREMINIQLREVLEADATDKM 148

## RESULT 10

C97308  
probable acetyltransferase [imported] - Clostridium acetobutylicum  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001

C:Accession: C97308  
R:Moiling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium  
A:Reference number: A96900; MWID:21359325; PMID:21359325

A:Accession: C97308  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-163 <NUR>  
A:Cross-references: GB:AE001437; PIDN:AAK81254.1; PID:g15026402; GSPDB:GNO0168  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC3322

Query Match 7.3%; Score 82; DB 2; Length 163;  
Best Local Similarity 24.5%; Pred. No. 4.2;  
Matches 37; Conservative 11; Mismatches 51; Indels 52; Gaps 6;

OY 82 EIERKEEDYTVTSFTFRNASSPIKYNTETVKAFFQYKNI 125  
DB 3 KISVKQMDYETAKQITKWNTEKPPSYNNSECEIKELLSGYFSAYDKNTYIGYCF 62  
OY 126 RNAIEYGVGGGL-----NITDPLFTDGLCDVFPVPNADG-----CELMVKSHY 172  
DB 63 GEAAGIPIVGQIGIYSKNTIDIGILNPLC-----GQGLGRDFECGL----- 107  
OY 173 KHVPDYCTFVFVNFPCAKD--RKYDIFNEECV 202  
DB 108 -----DFARNTLCAKDFRLTVATFNKPAI 131

## RESULT 11

T17996  
hypothetical protein A494R - Chlorella virus PBCV-1  
C:Species: Chlorella virus PBCV-1  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T17996  
R:Graves, M.V.; Van Etten, J.L.  
submitted to the EMBL Data Library, May 1999  
A:Reference number: Z18806  
A:Accession: T17996  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-360 <GRA>  
A:Cross-references: EMBL:U42580; NID:g4028886; PIDN:ANC96861.1  
A:Experimental source: specific host Chlorella strain NC64A  
C:Genetics:  
A:Note: A494R

Query Match 7.3%; Score 82; DB 2; Length 360;  
Best Local Similarity 25.8%; Pred. No. 12;  
Matches 41; Conservative 19; Mismatches 59; Indels 40; Gaps 8;

OY 35 HEELGK---YODAMKSIDQGVSVTVLAKTENDTGSWSGQFCLQVOEIERKEEDY 91  
DB 43 HEAVNGNIGLQFOGKKMKLENGV-----TKKTE-----KIKQIEDEINYL 84  
OY 92 VTSFTFRNASSPIKYNTETV-----KAVFOYKNIIR-NAIE---YQVGGGLNT 140  
DB 85 LDTI-----PFIEYDYKETSVDTEQNSVFOVSKNTHTNTRKYLFEHVEKVSNP 136  
OY 141 DTLFTDGLCDVFPVPNADGCELMVKSHYKHVPDYC 179  
DB 137 TLDAVTDRETVDOIYCTCGGOMELMVNSIQSDLVCEC 175

## RESULT 12

T18876  
hypothetical protein C03C10.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 11-Jan-2000

C:Accession: T18876  
R:Berts, M.  
submitted to the EMBL Data Library, August 1994  
A:Reference number: Z19036  
A:Accession: T18876





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2003, 17:13:11 ; Search time 40 Seconds

(without alignments)  
502.481 Million cell updates/sec

Title: DRET6

Perfect score: 1125

Sequence: 1 MKMOVLLITFVSALATQA.....DRKTYDIFNECVNGEPWL 209

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96.5	8.6	621	2	F97862 heat shock protein
2	96	8.5	650	2	G70169 heat shock protein
3	87	7.7	639	2	F71258 probable heat shock
4	86	7.6	2747	2	B49132 fat facets (fat) s
5	85	7.6	656	2	A87112 heat shock protein
6	84.5	7.5	621	2	H71645 heat shock protein
7	84.5	7.5	917	2	AF1394 autolysin, amidase
8	83	7.4	2285	2	T12076 probable transglyc
9	82.5	7.3	318	2	C69192 stomatin-like prot
10	82	7.3	163	2	C97308 probable acetyltra
11	82	7.3	360	2	T17996 hypothetical prote
12	82	7.3	381	2	T18766 hypothetical prote
13	82	7.3	390	2	D82922 conserved hypothet
14	82	7.3	637	2	T08530 trng protein - Ent
15	82	7.3	637	2	S22992 trng protein - Esc
16	81.5	7.2	627	2	D11729 dnaK-type molecula
17	80.5	7.2	621	2	A71961 chaperone and heat
18	79.5	7.1	621	2	B64546 90kDa chaperone -
19	79.5	7.1	725	2	AG2547 hypothetical prote
20	78.5	7.0	455	2	A75121 hypothetical prote
21	78.5	7.0	864	2	S60441 hypothetical prote
22	78	6.9	215	2	A26669 nodulin-22 precurs
23	78	6.9	523	1	O4CKA3 cytochrome P450 52
24	78	6.9	1871	2	S27938 hypothetical prote
25	77.5	6.9	223	2	T12076 ribonuclease (EC 3
26	77.5	6.9	506	2	T25870 hypothetical prote
27	77	6.8	241	2	T26909 hypothetical prote
28	77	6.8	331	2	D81409 probable periplasm
29	77	6.8	780	2	H84453 probable heat shoc

30	77	6.8	914	2	S70906 transferin-Dlindin
31	76.5	6.8	587	2	E87420 poly-beta-hydroxyb
32	76.5	6.8	635	2	T46407 probable RNA helic
33	76.5	6.8	807	2	T18454 hypothetical prote
34	76.5	6.8	2819	2	A90551 conserved hypothet
35	76	6.8	260	2	A55122 deoxycytidine kina
36	76	6.8	510	2	A71816 hypothetical prote
37	76	6.8	510	2	A64706 lipase-like protei
38	76	6.8	531	2	T18741 hypothetical prote
39	76	6.8	598	2	AG0890 hypothetical prote
40	76	6.8	781	2	S49340 probable arylsulfa
41	76	6.8	1031	2	T38411 heat-shock protein
42	76	6.8	1066	2	B95037 probable GTPase ac
43	76	6.8	1068	2	S73091 hyaluronidase [imp
44	75	6.8	1277	2	B45171 hypothetical prote
45	75.5	6.7	820	2	D81382 hypothetical prote
					probable transmemb

## ALIGNMENTS

## RESULT 1

F97862 heat shock protein htpg [imported] - Rickettsia conorii (strain Malish 7)

C:Species: Rickettsia conorii

C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 22-Oct-2001

C:Accession: F97862

R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.;

Science 293, 2093-2098, 2001

A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.

A:Reference number: A97700; MUID:21442074; PMID:11557893

A:Accession: F97862

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-621 <KUR>

A:Cross-references: GB:AE006914; PIDN:AAL03840.1; PID:915620441; GSPDB:GN00173

A:Genetics:

A:Gene: htpg

C:Superfamily: heat shock protein 90

Query Match	8.6%;	Score 96.5;	DB 2;	Length 621;
Best Local Similarity	21.0%;	Pred. No. 1.2;		
Matches	43;	Conservative	46;	Mismatches 81; Indels 35; Gaps 10;
QY	11	FVSALATQAETTSKAGENP--LMAHEILGKY--ODAMKSIOGVSVTYVLAKTYYEN 66		
DB	127	YSEFWADKVVTYSKAGEDVHTW-ESDGLGETVSDSDKEFTRGTEIVLHIKK---EE 182		
QY	67	DTGSMGSOFCLOVOEIERKEEDYTVTSVTFE-----RNASSPIKYYNVTETKAV 117		
DB	183	DT--FLDHFRR--LKHIVKYSVDHIAVPIFFDEAGNNEIQDLSAALMTPRKSEIPEQ 237		
QY	118	FQYGYKNTIRNAIE-----YQVGGGLNITPTLFTDGLCDVFRYPNADGCELEAYAK- 169		
DB	238	YKEFYKLSLYAIDDPWITMHNKNEGALIEFTMLFIPSSKTPDLPH-PDRKRRVRLYIKRV 296		
QY	170	----SHYKHPDVCYCFVNFPCARD 190		
DB	297	FISDENIDLPSYLFRLRGVDSDD 321		

## RESULT 2

G70169 heat shock protein 90 (htpg) homolog - Lyme disease spirochete

N:Alternate names: C62.5 heat shock protein

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 20-Aug-1999

C:Accession: G70169; I40247

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh

son, D.; Peterson, J.; Krelavag, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Yu

; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2003, 17:16:01 ; Search time 52 Seconds  
(without alignments)  
477.323 Million cell updates/sec

Title: DRET6  
Perfect score: 1125  
Sequence: 1 MKMVLLTFVSAALATQA.....DRKTYDIFNECVNGEPML 209

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:\*

```
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	282	25.1	171	14	US-10-085-572-7
2	282	25.1	171	14	US-10-087-195-7
3	251.5	22.4	182	14	US-10-085-572-8
4	251.5	22.4	182	14	US-10-087-195-8
5	239	21.2	172	14	US-10-085-572-6
6	239	21.2	172	14	US-10-087-195-6
7	102.5	9.1	722	9	US-09-815-242-10796
8	97.5	8.7	221	9	US-09-728-914-4
9	84	7.5	217	9	US-09-728-914-22
10	83	7.4	2285	10	US-09-932-183A-2
11	81	7.2	659	14	US-10-090-624-12
12	79.5	7.1	913	15	US-10-043-344-5
13	77	6.8	914	15	US-10-043-344-11
14	75	6.7	912	15	US-10-043-344-7
15	75	6.7	912	15	US-10-043-344-9

16	74.5	6.6	890	14	US-10-028-056-3	Sequence 3, Appl1
17	73	6.5	1417	10	US-09-753-143-78	Sequence 78, Appl
18	72.5	6.4	911	15	US-10-043-344-107	Sequence 107, Ap
19	72.5	6.4	2756	15	US-10-331-061-7	Sequence 7, Appl1
20	72	6.4	507	10	US-09-738-626-5458	Sequence 5458, Ap
21	72	6.4	956	14	US-10-121-032-63	Sequence 63, Appl
22	72	6.4	956	15	US-10-093-037-63	Sequence 63, Appl
23	71.5	6.4	564	15	US-10-081-872-120	Sequence 120, Ap
24	71.5	6.4	924	14	US-10-028-056-4	Sequence 4, Appl1
25	71.5	6.4	924	14	US-10-028-056-5	Sequence 5, Appl1
26	71.5	6.4	2012	10	US-09-808-602-68	Sequence 68, Appl
27	71.5	6.4	2012	11	US-09-801-198-57	Sequence 57, Appl
28	71	6.3	550	15	US-10-081-872-106	Sequence 106, Ap
29	71	6.3	618	15	US-10-295-403-166	Sequence 166, Ap
30	71	6.3	619	11	US-09-934-455-142	Sequence 142, Ap
31	71	6.3	677	10	US-09-862-027-38	Sequence 38, Appl
32	70.5	6.3	451	11	US-09-759-130B-413	Sequence 413, Ap
33	70.5	6.3	451	14	US-10-042-431-43	Sequence 43, Appl
34	70.5	6.3	470	11	US-09-759-130B-439	Sequence 439, Ap
35	70.5	6.3	470	14	US-10-042-431-69	Sequence 69, Appl
36	70	6.2	247	10	US-09-738-626-3674	Sequence 3674, Ap
37	70	6.2	576	9	US-09-731-175-4	Sequence 4, Appl1
38	70	6.2	576	15	US-10-011-348-30	Sequence 30, Appl
39	70	6.2	638	11	US-09-842-758-22	Sequence 22, Appl
40	70	6.2	853	15	US-10-156-761-8930	Sequence 8930, Ap
41	69.5	6.2	225	11	US-09-847-208-110	Sequence 110, Ap
42	69.5	6.2	373	11	US-09-759-1130B-376	Sequence 376, Ap
43	69.5	6.2	373	14	US-10-042-431-6	Sequence 6, Appl1
44	69.5	6.2	374	15	US-10-270-875-39	Sequence 39, Appl
45	69.5	6.2	374	15	US-10-270-878-39	Sequence 39, Appl

## ALIGNMENTS

RESULT 1									
US-10-085-572-7									
; Sequence 7, Application US/10085572									
; Publication No. US20020151499A1									
; GENERAL INFORMATION:									
; APPLICANT: Nuttall, Patricia, Ann									
; TITLE OF INVENTION: Treatment of Conjunctivitis									
; FILE REFERENCE: 2488-1-003									
; CURRENT APPLICATION NUMBER: US/10/085,572									
; CURRENT FILING DATE: 2002-02-27									
; PRIOR APPLICATION NUMBER: PCT/GB00/03282									
; PRIOR FILING DATE: 2000-08-24									
; PRIOR APPLICATION NUMBER: 9920674.0									
; PRIOR FILING DATE: 1999-09-01									
; NUMBER OF SEQ ID NOS: 8									
; SOFTWARE: FastSeq for Windows Version 4.0									
; SEQ ID NO 7									
; LENGTH: 171									
; TYPE: PRT									
; ORGANISM: FS-HBP 2									
US-10-085-572-7									
Query Match									
Best Local Similarity 25.1%; Score 282; DB 14; Length 171;									
Matches 66; Conservative 29; Mismatch 69; Indels 12; Gaps 8;									
QY	31	PLMAHEELGKQDADKSIDGVSYYVLAQTYENDTSGMSQSFCLQVLEERKEEDY	90						
DB	3	PMDADPANGAHODAKSLKADVENYVYKATYKNDP-VMGNDPFCVGMANDVDEDE	61						
QY	91	TYTSVETTF-RNASSPIKYYNVTETVKAFOYGYKRNRIAEVOVGGGLNTDPLTPTDGE	149						
DB	62	SIQAEELFLMNNADTNQF--ATEKTVAYVMYGC-NENMAFRYETEDGQFTDVIAYSD-D	117						
QY	150	LCDDVFYVPNAD---QCCELMVKKSHKHYPDVTCTFFVNFPCAKDRKRTYDIFNECV	202						
DB	118	NCDDVIVPCTDGNBEGYEIWT--TDYDNPANCLAKNFNEY-AVGRETRDVFYSACL	170						

## RESULT 2

US-10-087-195-7  
; Sequence 7, Application US/10087195  
; Publication No. US20020193306A1  
; GENERAL INFORMATION:  
; APPLICANT: Nuttall, Patricia, Anne  
; TITLE OF INVENTION: Treatment of Allergic Rhinitis  
; FILE REFERENCE: 2488-1-004  
; CURRENT APPLICATION NUMBER: US/10/087,195  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: PCT/GB00/03287  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 9920673.2  
; PRIOR FILING DATE: 1999-09-01  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 171  
; TYPE: PRT  
; ORGANISM: FS-HBP 2  
US-10-087-195-7

Query Match 25.1%; Score 282; DB 14; Length 171;  
Best Local Similarity 37.5%; Pred. No. 7, 3e-23;

Matches 66; Conservative 29; Mismatches 69; Indels 12; Gaps 8;

QY 31 PLMAHELLGKYDAMKSIDGVSVTVYLAKTTEYNDTSGMSQFCKLOV--QETIERKE 90  
DB 3 PDMADEANCAHODAMKSLKADVENYVYKATYKNDP-VGNDFTCVGMADVNEDEK 61  
QY 91 TVTSVFTE--RNASSPIKYYVETVKAPOYGYKKNIRNALEYOVGGSLNTDILIFDGE 149  
DB 62 SIAEPLFNNAADTNQF--ATEKVTAVKMYG--NRNARRYETEDQVETDVIAVSD-D 117  
QY 150 LCDVFPVFNAD--OGCELVWKKSHYKHPDYCTFFVNFCAKDRKTYDIFNEECV 202  
DB 118 NCOVIYVPGTGDGMEGEELMT--TDYDNIPANCLNFMNER-AVGRETRDVFISACL 170

## RESULT 3

US-10-085-572-8  
; Sequence 8, Application US/10085572  
; Publication No. US20020151499A1  
; GENERAL INFORMATION:  
; APPLICANT: Nuttall, Patricia, Ann  
; TITLE OF INVENTION: Treatment of Conjunctivitis  
; FILE REFERENCE: 2488-1-003  
; CURRENT APPLICATION NUMBER: US/10/085,572  
; CURRENT FILING DATE: 2002-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB00/03282  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 9920674.0  
; PRIOR FILING DATE: 1999-09-01  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 182  
; TYPE: PRT  
; ORGANISM: MS-HBP 1  
US-10-085-572-8

Query Match 22.4%; Score 251.5; DB 14; Length 182;  
Best Local Similarity 35.3%; Pred. No. 1, 7e-19;  
Matches 61; Conservative 29; Mismatches 66; Indels 17; Gaps 8;

QY 30 NPLMAHELLGKYDAMKSIDGVSVTVYLAKTTEYNDTSGMSQFCKLOV--QETIERKE 87  
DB 1 NPTWANEAKLGSYDAMKSLQDQONKRYLAQATQTTD-GVWGEEFTCVSVTAEKIGKK 59

QY 88 EDYTVTSVFTEFRNAS-SPIKYVNVETVKAFOYGYKKNIRNALEYOVGGSLNT--DTLI 144

DB 60 LNATI-----LYKNHLDLKESEHETITVMKAYDY---TTENGKIYETQGRTRQTFEEV 112

QY 145 FTDEGLCDVFPVNP--NADQG-CELVWKKSHYKHPDYCTFFVNFCAKDRKT 193

DB 113 FSDYKNCDVLFVPERKSGSDGDYELWVSEDKIDKIPDCKFTMAVFAQOQDEKT 165

## RESULT 4

US-10-087-195-8  
; Sequence 8, Application US/10087195  
; Publication No. US20020193306A1  
; GENERAL INFORMATION:  
; APPLICANT: Nuttall, Patricia, Anne  
; TITLE OF INVENTION: Treatment of Allergic Rhinitis  
; FILE REFERENCE: 2488-1-004  
; CURRENT APPLICATION NUMBER: US/10/087,195  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: PCT/GB00/03287  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 9920673.2  
; PRIOR FILING DATE: 1999-09-01  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 182  
; TYPE: PRT  
; ORGANISM: MS-HBP 1  
US-10-087-195-8

Query Match 22.4%; Score 251.5; DB 14; Length 182;  
Best Local Similarity 35.3%; Pred. No. 1, 7e-19;

Matches 61; Conservative 29; Mismatches 66; Indels 17; Gaps 8;

QY 30 NPLMAHELLGKYDAMKSIDGVSVTVYLAKTTEYNDTSGMSQFCKLOV--QETIERKE 87  
DB 1 NPTWANEAKLGSYDAMKSLQDQONKRYLAQATQTTD-GVWGEEFTCVSVTAEKIGKK 59  
QY 88 EDYTVTSVFTEFRNAS-SPIKYVNVETVKAFOYGYKKNIRNALEYOVGGSLNT--DTLI 144  
DB 60 LNATI-----LYKNHLDLKESEHETITVMKAYDY---TTENGKIYETQGRTRQTFEEV 112  
QY 145 FTDEGLCDVFPVNP--NADQG-CELVWKKSHYKHPDYCTFFVNFCAKDRKT 193  
DB 113 FSDYKNCDVLFVPERKSGSDGDYELWVSEDKIDKIPDCKFTMAVFAQOQDEKT 165

## RESULT 5

US-10-085-572-6  
; Sequence 6, Application US/10085572  
; Publication No. US20020151499A1  
; GENERAL INFORMATION:  
; APPLICANT: Nuttall, Patricia, Ann  
; TITLE OF INVENTION: Treatment of Conjunctivitis  
; FILE REFERENCE: 2488-1-003  
; CURRENT APPLICATION NUMBER: US/10/085,572  
; CURRENT FILING DATE: 2002-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB00/03282  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 9920674.0  
; PRIOR FILING DATE: 1999-09-01  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 172  
; TYPE: PRT  
; ORGANISM: FS-HBP1  
US-10-085-572-6

Query Match 21.2%; Score 239; DB 14; Length 172;

Query Match 21.2%; Score 239; DB 14; Length 172;

```

Best Local Similarity 34.5%, Pred. No. 3.7e-18;
Matches 61; Conservative 24; Mismatches 82; Indels 10; Gaps 7

OY      29 ENPLMAHEELLGKYODAMKSIDQGVSYTVYLAKTYYENDTGSWGSQFCLQVOETIERKEE 88
      : - - - - - : - - - - - : - - - - - : - - - - - : - - - - - :
Db      1 DKPWABEAAAGEHODAMKHLQKLEENYVDLIKATYKNDP--VWGNDFTCVGTAQNLNED 59
      : - - - - - : - - - - - : - - - - - : - - - - - : - - - - - :

OY      89 DYTTSVFTFENNASPIKYYVWTEYVKAVFQYGYKININAIYQVGGGLNTDILTFDGS 148
      : - - - - - : - - - - - : - - - - - : - - - - - : - - - - - :
Db      60 EKNVEAMFEMFNNAADTV-YQHTFEKATPDKKMIGY-NKENALITYQTEDQGLVDLVAFS- 116
      : - - - - - : - - - - - : - - - - - : - - - - - : - - - - - :

OY      149 ELCDVFVYPNAD--QGCELWVKRSHYKHVPDYCTFENVFCANKDRKTYDIFNECV 202
      : - - - - - : - - - - - : - - - - - : - - - - - : - - - - - :
Db      117 DNCYVYVALGPDGSGAGYELMA--TDYTDVPASCLEKRFNEYAA-GLPVRDYVTSQCL 170
      : - - - - - : - - - - - : - - - - - : - - - - - : - - - - - :

RESULT 6
US-10-087-195-6
; Sequence 6, Application US/10087195
; Publication No. US20020193306A1
; GENERAL INFORMATION:
; APPLICANT: Nuttall, Patricia, Anne
; APPLICANT: Paesen, Guido, Christiaan
; TITLE OF INVENTION: Treatment of Allergic Rhinitis
; FILE REFERENCE: 2488-1-004
; CURRENT APPLICATION NUMBER: US/10/087.195
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: PCT/GB00/03287
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 9920673.2
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 172
; TYPE: PRT
; ORGNISM: FS-HBPI
US-10-087-195-6

Query Match      21.2%, Score 239; DB 14; Length 172;
Best Local Similarity 34.5%, Pred. No. 3.7e-18;
Matches 61; Conservative 24; Mismatches 82; Indels 10; Gaps 7

OY      29 ENPLMAHEELLGKYODAMKSIDQGVSYTVYLAKTYYENDTGSWGSQFCLQVOETIERKEE 88
      : - - - - - : - - - - - : - - - - - : - - - - - : - - - - - :
Db      1 DKPWABEAAAGEHODAMKHLQKLEENYVDLIKATYKNDP--VWGNDFTCVGTAQNLNED 59
      : - - - - - : - - - - - : - - - - - : - - - - - : - - - - - :

OY      89 DYTTSVFTFENNASPIKYYVWTEYVKAVFQYGYKININAIYQVGGGLNTDILTFDGS 148
      : - - - - - : - - - - - : - - - - - : - - - - - : - - - - - :
Db      60 EKNVEAMFEMFNNAADTV-YQHTFEKATPDKKMIGY-NKENALITYQTEDQGLVDLVAFS- 116
      : - - - - - : - - - - - : - - - - - : - - - - - : - - - - - :

OY      149 ELCDVFVYPNAD--QGCELWVKRSHYKHVPDYCTFENVFCANKDRKTYDIFNECV 202
      : - - - - - : - - - - - : - - - - - : - - - - - : - - - - - :
Db      117 DNCYVYVALGPDGSGAGYELMA--TDYTDVPASCLEKRFNEYAA-GLPVRDYVTSQCL 170
      : - - - - - : - - - - - : - - - - - : - - - - - : - - - - - :

RESULT 7
US-09-815-242-10796
; Sequence 10796, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlson, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815.242

```

```

; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 10796
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10796

Query Match          9.1%; Score 102.5; DB 9; Length 722;
Best Local Similarity 26.9%; Pred. No. 0.023;
Matches 60; Conservative 26; Mismatches 72; Indels 65; Gaps 15;

QY 16 LATQAEITSA-----KAENELPMABEELCKYQDAWKSI-DQGVSVTVYVLAKTTY----- 64
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 204 LVGDRETSSTSSGFAVKKQNP-----ELIKRFNACLKMLKDNG--TYDKILNNLATGD 255
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 65 ENDIGSMGSOFCRCLQVOEIERKEEDYTVTSVFTFNNASSPIKYYNV-----TEFYKA 116
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 256 ETNQDAEQQK-----KITPKKEKYVIASDSTF-----APEFQNAQGDYGVIDVLYKR 306
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 117 V-----FQY-GYKIRNAIEI-QVGG---GLNITDITLTDTGELCDVFIYVNAQDG 162
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 307 AAELQGFTEVERKEFIQSSAYQAVESGADGNVAGMTITD-----DRKRAFDFSVYFDG 361
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 163 CELMWKSKH--YKHVPDYCTFEFVN-----FCADKRRTYD 195
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 362 IQIAVKKGNDKIKSYDDLKGGKVKYKIGTESADFLKKKKKYD 404
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
US-09-728-914-4
; Sequence 4, Application US/09728914
; Patent No. US20010046499A1
; GENERAL INFORMATION:
; APPLICANT: KANTOR, FRED S.
; APPLICANT: KANTOR, FRED S.
; APPLICANT: DAS, SUBRATA
; TITLE OF INVENTION: TICK ANTIGENS AND COMPOSITIONS AND METHODS COMPRISING
; FILE REFERENCE: YU-107
; CURRENT APPLICATION NUMBER: US/09/728,914
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/169,048
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: 60/240,716
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentlin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Ixodes scapularis
US-09-728-914-4

Query Match          8.7%; Score 97.5; DB 9; Length 221;
Best Local Similarity 23.4%; Pred. No. 0.015;
Matches 49; Conservative 43; Mismatches 80; Indels 37; Gaps 14;

```

```

OY      10 TVFAAALATQAE-----TSAGNPLMAHELLGKY--ODAMSIDOGSVFVLA 60
      11 : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB      31 TYSGTGTTTRGTCGARNITVTTAPBEDP-----SKTEONATRYVE--MATOWK 80
OY      61 KTYENDTGSWSQKCLQVOELERKEEDYTVTSV-PTFNASSPIKYVWTE--VKAV 117
      11 : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB      81 WRITD-VTDFSGNIVQCEENFRVMEKR---PTNYSQRYKSKN-SWEIIDEITLLIKDI 134
OY      118 FQYQKINRAIEYQVGGGLNTDTLL-FTDGEICDVFYYPNADQ---CELWKK-SHY 172
      11 : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB      135 GEHGFPMVMNRPPIC---IATDNLVLSNYVNCVLRIPFTNOGERHCDLMMANITLS 191
OY      173 KHVPCYCFVFNVCARDKRTYDIENEC 201
      11 : | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      192 QETPDCLNKFEEYC-NTQIYRVYYPSC 219

RESULT 9
US-09-728-914-22
; Sequence 22, Application US/09728914
; Patent No. US2001004649A1
; GENERAL INFORMATION:
; APPLICANT: KANTOR, FRED S.
; APPLICANT: FIKRIG, EROL
; APPLICANT: DAS, SUBRATA
; TITLE OF INVENTION: TICK ANTIGENS AND COMPOSITIONS AND METHODS COMPRISING
; TITLE OF INVENTION: THEM
; FILE REFERENCE: YU-107
; CURRENT APPLICATION NUMBER: US/09/728, 914
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/169, 048
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: 60/240, 716
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Ixodes scapularis
US-09-728-914-22

Query Match      7.5%; Score 84; DE 9; Length 217;
Best Local Similarity 24.1%; Pred. No. 0.45;
Matches 27; Conservative 17; Mismatches 48; Indels 20; Gaps 3;

OY      99 RNASSPIKYVNVETKAVFQYQKINRAIEYQVGGGLNTDTLLFTDS----- 148
      11 : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB      79 RNDSS-----YEFVKRTLSTISVGEQSPKRHTEIYEITGTSKSEKAVTKGYEEDYDNTY 134
OY      149 --ELCDVFYYPNAD---QGCELVWKKSHYKVPDYCTEFVNFCAKDRKTY 194
      11 : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB      135 ADTRCLILRISRTQKVPYLRSCLLMWKKTFLKPNLRHCRRLFDFVFCMMRRDFE 186

RESULT 10
US-09-932-183A-2
; Sequence 2, Application US/09932183A
; Patent No. US20020127641A1
; GENERAL INFORMATION:
; APPLICANT: Estell, David A.
; TITLE OF INVENTION: Proteases From Gram-Positive Organisms
; FILE REFERENCE: GC94C1-US
; CURRENT APPLICATION NUMBER: US/09/932,183A
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/308, 375
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/US98/18828
; PRIOR FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: EP9719636.4
; PRIOR FILING DATE: 1997-09-15
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0

```

```

; SEQ ID NO 2
; LENGTH: 2285
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-09-932-183A-2

Query Match          7.4%; Score 83; DB 10; Length 2285;
Best Local Similarity 23.3%; Pred. No. 15;
Matches 51; Conservative 27; Mismatches 79; Indels 62; Gaps 10

6 VLLTFVSAALATQAEITTSAKAGENPPLAHAEELLGKDYODAMKSID--OGSVTYVIAKTT 63
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
743 LLVSTLVGCAFAALG-----MAELISFPAEKAKANDFEQSQTINVAEITTT 790
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
64 YENDTGSWGSOFKLOVOEIERKEEDYTVSYTFERNMASSPIKYVNTETVKAFFQ---Y 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
791 NKDSTDKLIQOK-----ELQKVKRSRITS-----DEQDEYLVQVQLAQLAEPALVK 838
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 GY-----KIRNAIEYQVGGGLNTIDTLFTDGLCDVYVVPNADGCELMVKK 169
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
839 GYDSOGNAILKTNKELEKAIE-----NTKEYLAKKQETPDSAKKTFEDASKE--IKK 889
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
170 S-----HKHVPDYCTFEVFNFCADKRTYDIFNECYV 203
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
850 SKDELKQYKQIADYND-----KGRKRWLLIADDDY 920

RESULT 11
US-10-090-624-12
; Sequence 12, Application US/10090624
; Publication NO. US2002013235A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOTO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/10/090, 624
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Thermococcus celer
US-10-090-624-12

Query Match          7.2%; Score 81; DB 14; Length 659;
Best Local Similarity 25.9%; Pred. No. 4.5;
Matches 53; Conservative 20; Mismatches 62; Indels 70; Gaps 13

41 KYODAMK-----SIDGVSV--TVYIAKTTYENDTGSN--GSQFCLQVOEIERKEEDYT 91
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
436 KYDDYAKITLFGSAADKKSATHTEDVSGATETVATLIYDITSSDIDLTLIPNGNEVYS 495
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
92 TVSYTFERNMASSPIKYVNTET--TVKAVFOYGYKKIRNAIEYQV-----GGG- 136
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
496 YTAIVYGEK-----VGYVNPATGATWTVKVV---SYKGAAN---YQDVVYSDGSLSGSGGN 545
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
137 -----LNTITDLFTDGLCDVYVVPNADGCELMVKKSHKHVPDYCTFEVNFVC 187
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
546 PNPAPNPNPPTTTDQFET--GSVND--YMDTSD-----TFTMNVNS 583
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
188 AKDKRTYD-----FNEECYVNGEP 207
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
584 GATKITGDLTFDTSYNDLTLVDP 608
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```
RESULT 12
US-10-043-344-5
; Sequence 5, Application US/10043344
; Publication No. US20030088086A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Harkness, Robin E.
; APPLICANT: Schryvers, Anthony B.
; APPLICANT: Chong, Pele
; APPLICANT: Gray-Owen, Scott
; APPLICANT: Murdin, Andrew D.
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: TRANSFERRIN RECEPTOR GENES
; FILE REFERENCE: 1038-1221 MIS
; CURRENT APPLICATION NUMBER: US/10/043,344
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 08/649,518
; PRIOR FILING DATE: 1996-05-17
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 913
; TYPE: PRT
; ORGANISM: Haemophilus Influenzae
US-10-043-344-5
```

```
Query Match 7.1% Score 79.5; DB 15; Length 913;
Best Local Similarity 21.9%; Pred. No. 10;
Matches 48; Conservative 32; Mismatches 60; Indels 79; Gaps 13;

QY 11 FVSAALATQAEETTSKAGENPLMAHELLGKYOD-AKMS-IDGVSVTYVLAKTT 64
DB 593 YVDGLGIRYDVSNTKANEIYI-----SVGFKFNSNNTGIVIKPTMLDLSYRLS-TGF 646
QY 65 EN-----DTGSMGSOFCLOVOEIERKE-----DYTVTSVFERNASS 103
DB 647 RNPSFAEMYGMRGKNTDVIYIGFK-----PETS RNDEFGLALKGDFGNIISHSNA-- 700
QY 104 PIKYNTVETVKAFO-----YGYKNIRNAIEYOGGGLNTDTLFTDGLCDVFEY 155
DB 701 ---YRNLIAFAEELSKGTGKNGYGNMAONAKLV---GVNITAOIDFNG----- 745
QY 156 VPNAOCCELWVKSHYKHPD--YCTFEVNFVFCADRK 192
DB 746 -----LM-----KRIPYGWATFAFAYNRVAKVDOK 769
```

```
RESULT 13
US-10-043-344-11
; Sequence 11, Application US/10043344
; Publication No. US20030088086A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Harkness, Robin E.
; APPLICANT: Schryvers, Anthony B.
; APPLICANT: Chong, Pele
; APPLICANT: Gray-Owen, Scott
; APPLICANT: Murdin, Andrew D.
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: TRANSFERRIN RECEPTOR GENES
; FILE REFERENCE: 1038-1221 MIS
; CURRENT APPLICATION NUMBER: US/10/043,344
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 08/649,518
; PRIOR FILING DATE: 1996-05-17
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 914
; TYPE: PRT
; ORGANISM: Haemophilus Influenzae
US-10-043-344-11
```

```
Query Match 6.8% Score 77; DB 15; Length 914;
Best Local Similarity 23.0%; Pred. No. 20;
Matches 51; Conservative 24; Mismatches 61; Indels 86; Gaps 14;

QY 11 FVSAALATQAEETTSKAGENPLMAHELLGKYOD-AKMSIDGVSVTYVLAKTT----- 63
DB 595 YVDGLGIRYDVSNTKANEIYI-----SVGFKFNSN-----TGIVIKPTMLDLS 641
QY 64 YENDTG-----SMG-----SOFKCLVOEIERKE-----DYTVTSVF 96
DB 642 YRLSTGFRNPSFAEMYGMRGNGNSEYVIGFK-----PETS RNDEFGLALKGDFGNIIS 697
QY 97 TFRNASSPI-----KIYNTVETVKAFOYGYKNIRNAIEYOGGGLNTDTLFTDGLCD 152
DB 698 HFSNAYRNLIAFAEELSKGTGKA--NYGYHMAONAKLV---GVNITAOIDFNG----- 746
QY 153 VEVYPNAOCCELWVKSHYKHPD--YCTFEVNFVFCADRK 192
DB 747 -----LM-----KRIPYGWATFAFAYNRVAKVDOK 770
```

```
RESULT 14
US-10-043-344-7
; Sequence 7, Application US/10043344
; Publication No. US20030088086A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Harkness, Robin E.
; APPLICANT: Schryvers, Anthony B.
; APPLICANT: Chong, Pele
; APPLICANT: Gray-Owen, Scott
; APPLICANT: Murdin, Andrew D.
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: TRANSFERRIN RECEPTOR GENES
; FILE REFERENCE: 1038-1221 MIS
; CURRENT APPLICATION NUMBER: US/10/043,344
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 08/649,518
; PRIOR FILING DATE: 1996-05-17
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 912
; TYPE: PRT
; ORGANISM: Haemophilus Influenzae
US-10-043-344-7
```

```
Query Match 6.7% Score 75; DB 15; Length 912;
Best Local Similarity 21.6%; Pred. No. 32;
Matches 47; Conservative 33; Mismatches 60; Indels 78; Gaps 13;

QY 11 FVSAALATQAEETTSKAGENPLMAHELLGKYOD-AKMS-IDGVSVTYVLAKTT 64
DB 593 YVDGLGIRYDVSNTKANEIYI-----SVGFKFNSNNTGIVIKPTMLDLSYRLS-TGF 646
QY 65 ENDTGS-----WGSOFKCLVOEIERKE-----DYTVTSVFERNASS 103
DB 647 RNPSFSEMYGMRGKNDDEVYVIGFK-----PETS RNDEFGLALKGDFGNIISHSNA-- 700
QY 104 PIKYNTVETVKAFO-----YGYKNIRNAIEYOGGGLNTDTLFTDGLCDVFEY 156
DB 701 ---YRNLIAFAEELSKGTGKNGYGNMAONAKLV---GVNITAOIDFNG----- 744
QY 157 PNAOCCELWVKSHYKHPD--YCTFEVNFVFCADRK 192
DB 745 -----LM-----KRIPYGWATFAFAYNQVAKVDOK 768
```

```
RESULT 15
US-10-043-344-9
; Sequence 9, Application US/10043344
; Publication No. US20030088086A1
```





Db 1 KLAIIAVCFISTVSSVPSDITSEE-----EDLNTRKNDIMKTSK--SKVFWLRL 49  
QY 62 TYENDTSGWSQFCL-----OVDEI-----ERKEDYTVTSVF--TFRNASSP 104  
Db 50 RRTWVD-----GKRSICRYGVLRKDKANHIEIDIMSYDSQETIYKTKYITTRLCRSG 105  
QY 105 IKYNTVETVKAFOYQYKINRINAIEYOVGGLNTDTLFTDGLCDVFEVY----- 156  
Db 106 DRNH-----MGVSLQGYNH--TGIEYK-----MYYDDQCCALIKYTKDNRNQP 147  
QY 157 PNADGCELMVKKSHKHVPDY--CTFEVNFPCAKDRKTYD 195  
Db 148 QNLKACEMWATANDANSVNSIAACEVYORRCNPNNNSYDD 188

RESULT 2  
US-09-107-532A-5279  
; Sequence 5279, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD/ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Arinello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 5279:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 497 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (b) LOCATION 1...497  
; SEQUENCE DESCRIPTION: SEQ ID NO: 5279:  
US-09-107-532A-5279

Query Match 7.5%; Score 84.5; DB 4; Length 497;  
Best Local Similarity 22.6%; Pred. No. 1.2;  
Matches 47; Conservative 29; Mismatches 85; Indels 47; Gaps 9;

QY 26 KAGENPLMAHEELGKYODAKMSIDGCVSVTVYLAKTYYEND--TGSMSQKCLQVOYDI 83  
Db 15 KPEBGLMGATPAANOCBEGAMVDYDEKGLSVSDY--TFDSLPKPKKWTDDHMMTHOQV 71

QY 84 ERKEDYTVTSVTFEFNASSPIKYVNTETVKAVFQYKINRINAIEYOVGGLNTDTL 143  
Db 72 KANQDP--NSKYYKRRHNDYFHHFKEDIRLFAMGFQCYMSIAM-----TR 118  
QY 144 IF-----TDGELCDVFPYPNADQ-----GCELMVKKSHY-----KHVPD 177  
Db 119 IFPHGDEETPNEAGLFYDQVFEDECLKYGIETPVSLSHWEMPLYLVTEGGMFRRLIOF 178  
QY 178 YCTFEVNFPCAKDRKT--YDIENE-ECV 202  
Db 179 YRFAETVERKRYKKNKYKWMTEFNEINCV 206

RESULT 3  
US-09-308-375-2  
; Sequence 2, Application US/09308375  
; Patent No. 6300117  
; GENERAL INFORMATION:  
; APPLICANT: Genencor International, Inc.  
; TITLE OF INVENTION: Proteases From Gram-Positive Organisms  
; FILE REFERENCE: GC394-PCT  
; CURRENT APPLICATION NUMBER: US/09/308,375  
; CURRENT FILING DATE: 1999-05-14  
; EARLIER APPLICATION NUMBER: EP9719636.4  
; EARLIER FILING DATE: 1997-09-15  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 2285  
; TYPE: PRT  
; ORGANISM: Bacillus subtilis  
US-09-308-375-2

Query Match 7.4%; Score 83; DB 4; Length 2285;  
Best Local Similarity 23.3%; Pred. No. 16;  
Matches 51; Conservative 27; Mismatches 79; Indels 62; Gaps 10;

QY 6 VLLTFVSALATQAEETSAAKAGENPLMAHEELGKYODAKMSID--QGVSVTVYLAKT 63  
Db 743 LVTSTLVGAFALG-----WALESLISSFAKAKKAKDDFEGSQDTNVAITTT 790  
QY 64 YENDTSGWSQFCLQVOEIERKEDYTVTSVTFEFNASSPIKYVNTETVKAVFQ--Y 120  
Db 791 NKDSTDKLIQYK-----ELQVKRESRLTS-----DEQEYLVTOQLAOTFPALVK 838  
QY 121 GY-----KNIRNAIEYQVGGLNTDTLFTDGLCDVFPYPNADGCELMVKK 169  
Db 839 GYDSQGNALTKNKELEKAIE-----NTKVELALKKQETRDSAKRTFEDASKE--IRK 889  
QY 170 S-----HYKHVPDYCTFEVNFPCAKDRKTYDIFNEECVY 203  
Db 890 SKDELKQYKQIADYND-----KGRPKMDLIADDDY 920

RESULT 4  
US-08-894-818B-1  
; Sequence 1, Application US/08894818B  
; Patent No. 6261822  
; GENERAL INFORMATION:  
; APPLICANT: TAKAKURA, Hikaru  
; APPLICANT: MORISHITA, Mio  
; APPLICANT: YAMAMOTO, Katsuhiko  
; APPLICANT: MITTA, Masanori  
; APPLICANT: ASADA, Kiyozo  
; APPLICANT: TSUNASAWA, Susumu  
; APPLICANT: KATO, Ikinoshin  
; TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Brody and Nelmark  
; STREET: 419 Seventh Street N.W., Ste. 300  
; CITY: Washington



```
STATE: D.C.
COUNTRY: United States of America
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/894,818B
FILING DATE: 20-MAY-1998
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03253
FILING DATE: 07-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 323285/1995
FILING DATE: 12-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TAKAKURA-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 659 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-894-818B-1
```

```
Query Match 7.28; Score 81; DB 3; Length 659;
Best Local Similarity 25.98; Pred. No. 4.2;
Matches 53; Conservative 20; Mismatches 62; Indels 70; Gaps 13;
```

```
QY 41 KYDAMK-----SIDGVSV-TYVLAKTTENDTGSW--GSQKCLQVDEIERKEDYT 91
DB 436 KYDYAKLTFTGSVADGASATHTPDVSGATFTATLTLYWDGSSDIDLTLDPNGNEVDYS 495
QY 92 VTSVETFRNASSPIKYNVTE---TKAVFQYGYKNIRNAIEYOV-----GGG- 136
DB 496 YTAIYGEK---VGIYNPTAGTWTVKYV---STKGAAN---YQVDVYSDLSQSGGNGN 545
QY 137 -----LINTDTLFTDGLCDVFPVPNADGCELMVKSHYKHVPDYCTEFVNFYC 187
DB 546 PNPMPNPNPTPTDTQFTT-GSVND--YMDTSD-----FTTMNVNS 583
QY 188 AKDRKTYDI-----FNEECYNGEP 207
DB 584 GATKITGDLTFDTSYNDLTLYDP 608
```

```
RESULT 5
US-08-894-818B-5
Sequence 5, Application US/08894818B
Patent No. 6261822
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: YAMAMOTO, Katsuhiko
APPLICANT: WITTA, Masanori
APPLICANT: ASADA, Kiyozo
APPLICANT: TSUNASAWA, Susumu
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Browdy and Nelmark
STREET: 419 Seventh Street N.W., Ste. 300
CITY: Washington
```

```
STATE: D.C.
COUNTRY: United States of America
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/894,818B
FILING DATE: 20-MAY-1998
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03253
FILING DATE: 07-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 323285/1995
FILING DATE: 12-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TAKAKURA-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 659 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-894-818B-5
```

```
Query Match 7.28; Score 81; DB 3; Length 659;
Best Local Similarity 25.98; Pred. No. 4.2;
Matches 53; Conservative 20; Mismatches 62; Indels 70; Gaps 13;
```

```
QY 41 KYDAMK-----SIDGVSV-TYVLAKTTENDTGSW--GSQKCLQVDEIERKEDYT 91
DB 436 KYDYAKLTFTGSVADGASATHTPDVSGATFTATLTLYWDGSSDIDLTLDPNGNEVDYS 495
QY 92 VTSVETFRNASSPIKYNVTE---TKAVFQYGYKNIRNAIEYOV-----GGG- 136
DB 496 YTAIYGEK---VGIYNPTAGTWTVKYV---STKGAAN---YQVDVYSDLSQSGGNGN 545
QY 137 -----LINTDTLFTDGLCDVFPVPNADGCELMVKSHYKHVPDYCTEFVNFYC 187
DB 546 PNPMPNPNPTPTDTQFTT-GSVND--YMDTSD-----FTTMNVNS 583
QY 188 AKDRKTYDI-----FNEECYNGEP 207
DB 584 GATKITGDLTFDTSYNDLTLYDP 608
```

```
RESULT 6
US-09-445-472-12
Sequence 12, Application US/09445472
Patent No. 6358726
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: SHIMOTO, Tomoko
APPLICANT: KATO, Kiyozo
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
FILE REFERENCE: TAKAKURA-6
CURRENT APPLICATION NUMBER: US/09/445,472
CURRENT FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 151969/1997
PRIOR FILING DATE: 1997-06-10
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.0
```

SEQ ID NO 12  
LENGTH: 659  
TYPE: PRF  
ORGANISM: Thermococcus celer  
US-09-445-472-12

Query Match 7.2%; Score 81; DB 4; Length 659;  
Best Local Similarity 25.9%; Pred. No. 4.2;  
Matches 53; Conservative 20; Mismatches 62; Indels 70; Gaps 13;

QY 41 KYODAK-----SIOGVSV--TYVAKTYENDTSGW--GSOFKCLOVEIERKEDYT 91  
DB 436 KIDVYAKLFTGTVADKSGATHTFDVSGATFVATLYWDTGSSDIDLXLDPNGNEVDY 495  
QY 92 VTSVETFRNASPIKRYNTE--TVKAVFOYGYKNIRNAIEYOV-----GGG- 136  
DB 496 YTAHYGEK---VGYNPTAGTWYKVV---SYKGAN---YQVDVDSGSLSSGGGN 545  
QY 137 -----LNTDPLFTDGLCDVFFYPNADGCELMVKKSHYKHPDCTFEVNFVC 187  
DB 546 PNEPNPNPTPTDTOTFT-GSYND--YMDTSD-----FTMNVNS 583  
QY 188 AKDKTYDI-----FNEECYNGEP 207  
DB 584 GATKITGDLFTDTSYNDLILYDP 608

## RESULT 7

US-08-487-890A-5  
Sequence 5, Application US/08487890A  
Patent No. 5708149

## GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena  
APPLICANT: Harkness, Robin  
APPLICANT: Schryvers, Anthony  
APPLICANT: Chong, Pele  
APPLICANT: Gray-Owen, Scott  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Murdin, Andrew  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Transferrin Receptor Genes  
NUMBER OF SEQUENCES: 147  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,890A  
FILING DATE: 07-JUN-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/175,116  
FILING DATE: 29-DEC-1993

PRIOR APPLICATION NUMBER: US 08/148,968  
FILING DATE: 08-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-466 MIS:jb

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:

LENGTH: 913 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-487-890A-5

Query Match 7.1%; Score 79.5; DB 1; Length 913;  
Best Local Similarity 21.9%; Pred. No. 9.9;  
Matches 48; Conservative 32; Mismatches 60; Indels 79; Gaps 13;

QY 11 FVSAALATQAEITTSARAGENPLMAHEELCKYOD-AWKS-----IDQGVSVTVLAKTIT 64  
DB 593 YVDLGLGMRVDSRTNANESTI---SVGKFNFSNITGVIKPTFEMLDLSTRLS-TGR 646  
QY 65 EN-----DTSGMSGOSFKCLOVEIERKE-----DYTVSVTFNRASS 103  
DB 647 RNSPFAEMTGRGKADTDVITGFK---PETSRRDFELAKGDFGNIEHSFNA--700  
QY 104 PIKYVNTETVKAVFQ-----YGYKNIRNAIEYOVGGGLNITDPLFTDGLCDVFX 155  
DB 701 ---YRNLIAPAEELSKNGTGGKNGYGHNNQNAKLV---GVNITQOLDENG-----745  
QY 156 VPNAOGCELMVKKSHYKHPD--YCTFEVNFVNCADRK 192  
DB 746 -----LW-----KRIPYGYATFAYNRVYKVDOK 769

## RESULT 8

US-08-478-435-5  
Sequence 5, Application US/08478435  
Patent No. 5922323

## GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena  
APPLICANT: Harkness, Robin  
APPLICANT: Schryvers, Anthony  
APPLICANT: Chong, Pele  
APPLICANT: Gray-Owen, Scott  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Murdin, Andrew  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Transferrin Receptor Genes  
NUMBER OF SEQUENCES: 147  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/478,435  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/337,483  
FILING DATE: 08-NOV-1994

PRIOR APPLICATION NUMBER: US 08/175,116  
FILING DATE: 29-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-462 MIS:vg

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:

```

: TELEPHONE: (416) 595-1155
:
: TELEFAX: (416) 595-1163
:
: INFORMATION FOR SEQ ID NO: 5:
:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 913 amino acids
:
: TYPE: amino acid
:
: STRANDEDNESS: single
:
: TOPOLOGY: linear
:
US-08-478-435-5

```

Query Match	7.1%	Score 79.5	DB 2	Length 913
Best Local Similarity	21.9%	Pred No. 9.9		
Matches 48	Conservative 32	Mismatches 60	Indels 79	Gaps 13

[illegible]

RESULT 9  
 US-08-1337-483--5  
 Sequence 5, Application US/08337483  
 Patent No. 5922562  
 GENERAL INFORMATION:  
 APPLICANT: Loosmore, Sheena  
 APPLICANT: Harkness, Robin  
 APPLICANT: Schryvers, Anthony  
 APPLICANT: Chong, Pele  
 APPLICANT: Gray-Owen, Scott  
 APPLICANT: Yang, Yan-ping  
 APPLICANT: Murlin, Andrew  
 APPLICANT: Klein, Michel  
 TITLE OF INVENTION: Transferrin Receptor Genes  
 NUMBER OF SEQUENCES: 147  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Slim & McBurney  
 STREET: Suite 701, 330 University Avenue  
 CITY: Toronto  
 STATE: Ontario  
 COUNTRY: Canada  
 ZIP: M5G 1R7  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.2.22-1994  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/337,483  
 FILING DATE: 08-NOV-1994  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Stewart, Michael I  
 REGISTRATION NUMBER: 24,973  
 REFERENCE/DOCKET NUMBER: 1038-410 MIS:jb  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (416) 595-1155  
 TELEFAX: (416) 595-1163  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 913 amino acids  
 TYPE: amino acid

```

; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-337-483-5

```

Query Match	7.1%	Score 79.5	DB 2	length 913
Best Local Similarity	21.9%	Pred. No. 9.9		
Matches 48	Conservative 32	Mismatches 60	Indels 79	Gaps 13

RESULT 10  
US-08-478-373-5  
; Sequence 5, Application US/08478373  
; Patent No. 5922841

```

1  APPLICANT: Harkness, Robin
2  APPLICANT: Schryvers, Anthony
3  APPLICANT: Chong, Pele
4  APPLICANT: Gray-Owen, Scott
5  APPLICANT: Yang, Yan-Ping
6  APPLICANT: Murdin, Andrew
7  APPLICANT: Klein, Michel
8  TITLE OF INVENTION: Transferrin Receptor Genes
9  NUMBER OF SEQUENCES: 147
10 CORRESPONDENCE ADDRESS:
11 ADDRESSEE: Slim & McBurney
12 STREET: Suite 701, 530 University Avenue
13 CITY: Toronto
14 STATE: Ontario
15 COUNTRY: Canada
16 ZIP: M5G 1R7
17
18 COMPUTER READABLE FORM:
19 MEDIUM TYPE: Floppy disk
20 COMPUTER: IBM PC compatible
21 OPERATING SYSTEM: PC-DOS/MS-DOS
22 SOFTWARE: Patent Release #1.0, Version #1.25
23 CURRENT APPLICATION DATA:
24 APPLICATION NUMBER: US/08/478,373
25 FILING DATE: 07-JUN-1995
26 CLASSIFICATION: 435
27 PRIOR APPLICATION DATA:
28 APPLICATION NUMBER: US 08/337,483
29 FILING DATE: 08-NOV-1994
30 CLASSIFICATION: 435
31 PRIOR APPLICATION DATA:
32 APPLICATION NUMBER: US 08/175,116
33 FILING DATE: 29-DEC-1993
34 PRIOR APPLICATION DATA:
35 APPLICATION NUMBER: US 08/148,968
36 FILING DATE: 08-NOV-1993
37 ATTORNEY/AGENT INFORMATION:
38 NAME: Stewart, Michael I
39 REGISTRATION NUMBER: 24,973
40 REFERENCE/DOCKET NUMBER: 1038-463 MTS:v9
41 TELECOMMUNICATION INFORMATION:
42 TELEPHONE: (416) 595-1155
43 TELEFAX: (416) 595-1163

```

RESULT 9  
 US-08-1337-483--5  
 Sequence 5, Application US/08337483  
 Patent No. 5922562  
 GENERAL INFORMATION:  
 APPLICANT: Loosmore, Sheena  
 APPLICANT: Harkness, Robin  
 APPLICANT: Schryvers, Anthony  
 APPLICANT: Chong, Pele  
 APPLICANT: Gray-Owen, Scott  
 APPLICANT: Yang, Yan-ping  
 APPLICANT: Murlin, Andrew  
 APPLICANT: Klein, Michel  
 TITLE OF INVENTION: Transferrin Receptor Genes  
 NUMBER OF SEQUENCES: 147  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Slim & McBurney  
 STREET: Suite 701, 330 University Avenue  
 CITY: Toronto  
 STATE: Ontario  
 COUNTRY: Canada  
 ZIP: M5G 1R7  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.2.22-1994  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/337,483  
 FILING DATE: 08-NOV-1994  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Stewart, Michael I  
 REGISTRATION NUMBER: 24,973  
 REFERENCE/DOCKET NUMBER: 1038-410 MIS:jb  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (416) 595-1155  
 TELEFAX: (416) 595-1163  
 INFORMATION FOR SEQ. ID NO.: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 913 amino acids  
 TYPE: amino acid

INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 913 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

US-08-478-373-5

Query Match 7.1%; Score 79.5; DB 2; Length 913;  
Best Local Similarity 21.9%; Pred. No. 9.9;  
Matches 48; Conservative 32; Mismatches 60; Indels 79; Gaps 13;

QY 11 FVSAALATQAEITSAKAGNPILNAHELLKQYD-AWKS-----IDQGVSVTVLAKTY 64  
DB 593 YVLDGLGMRYDVSRNKANESTI-----SVGKFKNFNWTGIVIKPTFMDLSYRLS-TGF 646  
QY 65 EN-----DTGSMGSOFKLOVQEIETKE-----DYTVSVTFPRNASS 103  
DB 647 RNPSFAMTGMWRGKDTDYIGKFK-----PSTRNDFGLALKGDFGNIETSHFSNA-- 700  
QY 104 PIKYVNTETVKAFFQ-----YGYKNIRNAIEYQVGGILNTDTLFTDGLCDVY 155  
DB 701 --YRNLIAPAEELSKNGTGGKNGYGHNNQNKLV-----GVNITQDLDFNG----- 745  
QY 156 VPNADGCELWVKSHKHVPD--YCTFVNVFCANDRK 192  
DB 746 -----LM-----KRIPYGMVAFVAYNRVXVDOK 769

RESULT 11

US-08-474-671-5  
Sequence 5, Application US/08474671  
Patent No. 6008326

GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena  
APPLICANT: Harkness, Robin  
APPLICANT: Schryvers, Anthony  
APPLICANT: Chong, Pele  
APPLICANT: Gray-Owen, Scott  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Murdin, Andrew  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Transferrin Receptor Genes  
NUMBER OF SEQUENCES: 147  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/474,671

FILING DATE: 07-JUN-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/337,483

FILING DATE: 08-NOV-1994

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/175,116

FILING DATE: 29-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/148,968

FILING DATE: 08-NOV-1993

ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael I

REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-465 MIS:V9  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 913 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-474-671-5

Query Match 7.1%; Score 79.5; DB 3; Length 913;  
Best Local Similarity 21.9%; Pred. No. 9.9;  
Matches 48; Conservative 32; Mismatches 60; Indels 79; Gaps 13;

QY 11 FVSAALATQAEITSAKAGNPILNAHELLKQYD-AWKS-----IDQGVSVTVLAKTY 64  
DB 593 YVLDGLGMRYDVSRNKANESTI-----SVGKFKNFNWTGIVIKPTFMDLSYRLS-TGF 646  
QY 65 EN-----DTGSMGSOFKLOVQEIETKE-----DYTVSVTFPRNASS 103  
DB 647 RNPSFAMTGMWRGKDTDYIGKFK-----PSTRNDFGLALKGDFGNIETSHFSNA-- 700  
QY 104 PIKYVNTETVKAFFQ-----YGYKNIRNAIEYQVGGILNTDTLFTDGLCDVY 155  
DB 701 --YRNLIAPAEELSKNGTGGKNGYGHNNQNKLV-----GVNITQDLDFNG----- 745  
QY 156 VPNADGCELWVKSHKHVPD--YCTFVNVFCANDRK 192  
DB 746 -----LM-----KRIPYGMVAFVAYNRVXVDOK 769

RESULT 12

US-08-483-577A-5  
Sequence 5, Application US/08483577A  
Patent No. 6015688

GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena  
APPLICANT: Harkness, Robin  
APPLICANT: Schryvers, Anthony  
APPLICANT: Chong, Pele  
APPLICANT: Gray-Owen, Scott  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Murdin, Andrew  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Transferrin Receptor Genes  
NUMBER OF SEQUENCES: 160  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/483,577A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/337,483

FILING DATE: 08-NOV-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/175,116

FILING DATE: 29-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/148,968

FILING DATE: 08-NOV-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-511  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 913 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-483-577A-5

Query Match 7.18; Score 79.5; DB 3; Length 913;  
Best Local Similarity 21.98; Pred. No. 9.9;  
Matches 48; Conservative 32; Mismatches 60; Indels 79; Gaps 13;

OY 11 FVSAALTAQAEETTSKAGENPLMAHEELGKYOD-AWKS-----IDQGVSTVYLAQTTY 64  
DB 593 YVDGLGMRDYVSTKANESTI-----SVGKFKFNSMTGIVIPTEMLDLYRLS-TGF 646  
OY 65 EN-----DTGWSGQFKCLOVEIERKEE-----DYTVTSVTFPRNASS 103  
DB 647 RNPFAEMYGWRGCKDIDYIGKFK-----PETSNDQFGLALGDFGNIEISHPSNA-- 700  
OY 104 PIKYNTETEVKAVQ-----YQKNIRNALEYOVGGGLNTDILIFDGLCDVYF 155  
DB 701 ---YRNLIAPAEELSKNGTGGKNGYGNMAQNAKLV---GVNITADLFNG----- 745  
OY 156 VPNAOCCELWVKSHYKHPD--YCTFEVNFVFCADRK 192  
DB 746 -----LM-----KRIPYGMTATFAIRKRVKDDK 769

RESULT 13  
US-08-897-438-5  
Sequence 5, Application US/080897438  
Patent No. 6262016  
GENERAL INFORMATION:  
APPLICANT: Loosmore, Sheena  
APPLICANT: Harkness, Robin  
APPLICANT: Schryvers, Anthony  
APPLICANT: Chong, Pele  
APPLICANT: Gray-Owen, Scott  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Murdin, Andrew  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Transferrin Receptor Genes  
NUMBER OF SEQUENCES: 160  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/897,438  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/483,577  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/337,483  
FILING DATE: 08-NOV-1994

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/175,116  
FILING DATE: 29-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,968  
FILING DATE: 08-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-720  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 913 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-897-438-5

Query Match 7.18; Score 79.5; DB 3; Length 913;  
Best Local Similarity 21.98; Pred. No. 9.9;  
Matches 48; Conservative 32; Mismatches 60; Indels 79; Gaps 13;

OY 11 FVSAALTAQAEETTSKAGENPLMAHEELGKYOD-AWKS-----IDQGVSTVYLAQTTY 64  
DB 593 YVDGLGMRDYVSTKANESTI-----SVGKFKFNSMTGIVIPTEMLDLYRLS-TGF 646  
OY 65 EN-----DTGWSGQFKCLOVEIERKEE-----DYTVTSVTFPRNASS 103  
DB 647 RNPFAEMYGWRGCKDIDYIGKFK-----PETSNDQFGLALGDFGNIEISHPSNA-- 700  
OY 104 PIKYNTETEVKAVQ-----YQKNIRNALEYOVGGGLNTDILIFDGLCDVYF 155  
DB 701 ---YRNLIAPAEELSKNGTGGKNGYGNMAQNAKLV---GVNITADLFNG----- 745  
OY 156 VPNAOCCELWVKSHYKHPD--YCTFEVNFVFCADRK 192  
DB 746 -----LM-----KRIPYGMTATFAIRKRVKDDK 769

RESULT 14  
US-08-637-654-5  
Sequence 5, Application US/08637654  
Patent No. 6358727  
GENERAL INFORMATION:  
APPLICANT: Loosmore, Sheena M  
APPLICANT: Harkness, Robin E  
APPLICANT: Schryvers, Anthony B  
APPLICANT: Chong, Pele  
APPLICANT: Gray-Owen, Scott  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Murdin, Andrew D  
APPLICANT: Klein, Michel H  
TITLE OF INVENTION: HAEMOPHILUS TRANSFERRIN RECEPTOR GENES  
NUMBER OF SEQUENCES: 147  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/637,654  
FILING DATE: 05-AUG-1996  
CLASSIFICATION: 435





XX WPI: 1998-018506/02.  
 DR N-PSDB; AAV00230.  
 XX  
 XX New vasoactive amine binding proteins and related nucleic acid,  
 PT vectors - transformed cells and transgenic animals, used for  
 PT assaying or removing histamine and as antihistamine or  
 PT anti-inflammatory agents

Example 2; Fig 4; 44pp; English.

XX This protein comprises tick Dermacentor reticulatus (Dr) novel  
 CC vasoactive amine binding protein (VABP) D.RET6. Its amino acid  
 CC sequence was deduced from a cDNA clone (see AAV00230) obtained  
 CC from a Dr salivary gland cDNA library. 3 Novel VABPs, designated  
 CC FS-HBP1, FS-HBP2 and MS-HBP1 (see AAV37446-48), of the tick  
 CC Rhipicephalus appendiculatus have also been identified. The  
 CC VABPs can be expressed in host cells using e.g. a baculovirus  
 CC expression system. They can be used: (i) to assay histamine (or  
 CC other VA such as serotonin) in body fluids or cell culture  
 CC supernatants, e.g. to monitor the effect of allergens; (ii) for  
 CC binding VA, e.g. to remove histamine from blood, food, cell  
 CC cultures etc.; (iii) as an antihistamine or anti-inflammatory  
 CC agents, e.g. for treating insect, snake or scorpion bites or  
 CC dermatitis, or as a carrier for slow release of histamine-related  
 CC compounds; (iv) in vaccines to protect against metazoan parasites,  
 CC especially in animals; (v) as reagents for studying inflammation,  
 CC involvement of VA in ulcer formation or the immune response etc.  
 CC VABPs provide a more sensitive assay for histamine than  
 CC low-affinity antibodies currently used. They may also be more  
 CC effective and safer than conventional antihistamines.

XX Sequence 209 AA;

Query Match 100.0%; Score 1125; DB 19; Length 209;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-104;  
 Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKMOVLLTLFFVSAALTAQAEFTTSKAGENPLMAHEBLCKKYDAMKSIDGVSYYVLA 60  
 DB 1 MKMOVLLTLFFVSAALTAQAEFTTSKAGENPLMAHEBLCKKYDAMKSIDGVSYYVLA 60  
 QY KTTENTGSGMSQFKLOVEIERKEDYTVTSVFERNASSPIKYYNTETVKAFFQY 120  
 DB 61 KTTENTGSGMSQFKLOVEIERKEDYTVTSVFERNASSPIKYYNTETVKAFFQY 120  
 QY 121 GYKNIRNAIEYQVGGGINTDFTLFTDGLCDVFFVYNADGCELMWKKSHYKHVPDYCT 180  
 DB 121 GYKNIRNAIEYQVGGGINTDFTLFTDGLCDVFFVYNADGCELMWKKSHYKHVPDYCT 180  
 QY 181 FVENVFCADKDKTYDIFNECVYNGEPWL 209  
 DB 181 FVENVFCADKDKTYDIFNECVYNGEPWL 209

RESULT 2

AA18081  
 ID AAY18081 standard; Protein; 209 AA.

XX AAY18081;

DT 06-AUG-1999 (first entry)

XX Histamine binding protein D.RET6..

KW Histamine binding protein; serotonin binding compound; inflammation;  
 KW gastric acid secretion; allergy; type I hypersensitivity reaction;

KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
 KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
 KW respiratory disease; coronary heart disease; cellular growth regulator;  
 KW tissue repair; blood-sucking ectoparasite; therapy.

XX Rhipicephalus appendiculatus.  
 OS

XX WO9927104-A1.

XX 03-JUN-1999.

XX 26-NOV-1998; 98WO-GB03530.

XX 26-JUN-1998; 98GB-0013917.

XX 26-NOV-1997; 97GB-0025046.

XX (OXFO-) OXFORD VACS LTD.

XX Nuttall PA, Paesen GC;

XX WPI: 1999-357841/30.

XX N-PSDB; AAX76967.

PT Histamine and serotonin binding compounds useful for the treatment  
 PT of allergies  
 PS Claim 1; Fig 4; 84pp; English.

XX This sequence is an example of a histamine or serotonin binding  
 CC compound (A), of the invention. The compounds are useful for regulating  
 CC the action of histamine and serotonin (in e.g. inflammation and gastric  
 CC acid secretion), the detection, quantification and removal of histamine  
 CC or serotonin (in animals, plants, cell cultures, food materials, or  
 CC humans) and in the treatment of various diseases and allergies  
 CC (e.g. type I hypersensitivity reactions, urticaria, asthma, allergic  
 CC rhinitis (hay fever), atopic dermatitis, insect bites and food and drug  
 CC allergies, abnormal blood pressure, migraine, psychological disorders,  
 CC respiratory disease, and coronary heart disease). Histamine may also be  
 CC used to regulate cellular growth and tissue repair. The molecules may  
 CC also be used as components of vaccines directed against blood-sucking  
 CC ectoparasites.

XX Sequence 209 AA;

Query Match 100.0%; Score 1125; DB 20; Length 209;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-104;  
 Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKMOVLLTLFFVSAALTAQAEFTTSKAGENPLMAHEBLCKKYDAMKSIDGVSYYVLA 60  
 DB 1 MKMOVLLTLFFVSAALTAQAEFTTSKAGENPLMAHEBLCKKYDAMKSIDGVSYYVLA 60  
 QY KTTENTGSGMSQFKLOVEIERKEDYTVTSVFERNASSPIKYYNTETVKAFFQY 120  
 DB 61 KTTENTGSGMSQFKLOVEIERKEDYTVTSVFERNASSPIKYYNTETVKAFFQY 120  
 QY 121 GYKNIRNAIEYQVGGGINTDFTLFTDGLCDVFFVYNADGCELMWKKSHYKHVPDYCT 180  
 DB 121 GYKNIRNAIEYQVGGGINTDFTLFTDGLCDVFFVYNADGCELMWKKSHYKHVPDYCT 180  
 QY 181 FVENVFCADKDKTYDIFNECVYNGEPWL 209  
 DB 181 FVENVFCADKDKTYDIFNECVYNGEPWL 209

RESULT 3

AA18086  
 ID AAY18086 standard; Protein; 203 AA.

XX AAY18086;

DT 06-AUG-1999 (first entry)

XX Histamine binding protein Ih/Bm-HBP2.

KW Histamine binding protein; serotonin binding compound; inflammation;  
 KW gastric acid secretion; allergy; type I hypersensitivity reaction;

KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
 KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;



KW respiratory disease; coronary heart disease; cellular growth regulator;  
 KW tissue repair; blood-sucking ectoparasite; therapy.  
 XX  
 OS Boophilus microplus.  
 OS Ixodes hexagonus.  
 OS Synthetic.  
 PN WO9927104-A1.  
 XX  
 PD 03-JUN-1999.  
 XX  
 XX 26-NOV-1998; 98WO-GB03530.  
 XX  
 PR 26-JUN-1998; 98GB-0013917.  
 PR 26-NOV-1997; 97GB-0025046.  
 XX  
 PA (OXFO-) OXFORD VACS LTD.  
 XX  
 PI Nuttall PA, Paesen GC;  
 DR WPI: 1999-357841/30.  
 DR N-PSDB; AAX76971.  
 XX  
 PT Histamine and serotonin binding compounds useful for the treatment  
 of allergies  
 XX  
 PS Claim 14; Fig 8; 84pp; English.  
 CC This sequence is an example of a histamine or serotonin binding  
 compound (A), of the invention. cDNA encoding this sequence was isolated  
 from a mixed Boophilus microplus/Ixodes hexagonus cDNA expression  
 CC library. The compounds are useful for regulating the action of histamine  
 and serotonin (in e.g. inflammation and gastric acid secretion), the  
 CC detection, quantification and removal of histamine or serotonin (in  
 CC animals, plants, cell cultures, food materials, or humans) and in the  
 CC treatment of various diseases and allergies (e.g. type I hypersensitivity  
 CC reactions, urticaria, asthma, allergic rhinitis (hay fever), atopic  
 CC dermatitis, insect bites and food and drug allergies, abnormal blood  
 CC pressure, migraine, psychological disorders, respiratory disease, and  
 CC coronary heart disease). Histamine may also be used to regulate cellular  
 CC growth and tissue repair. The molecules may also be used as components of  
 CC vaccines directed against blood-sucking ectoparasites.  
 XX  
 SQ Sequence 203 AA:  
 Query Match 32.4%; Score 364; DB 20; Length 203;  
 Best Local Similarity 38.1%; Pred. No. 4e-28;  
 Matches 75; Conservative 38; Mismatches 78; Indels 6; Gaps 5;

QY 6 VLLTFVSAALATQAEETTSKAGENRPLMAHEBLCKYODAMKSIDOGSVTVYLAKTYYE 65  
 DB 8 VYLLAVTAADAPSSSTRNEPLEKTYW-HNOTLGRYODAMKSIDOGSVTVYLAKTYYE 66  
 QY 66 NDTGSMGQFKLOVQEIERKEEDYTVSVTFERNASSPIKYNTVETVKAVFOYGYKNI 125  
 DB 67 NDS-VWGNFTCLSTVTSKYSTFTVEYNTTYKNQSQ--QWVSSENVTAQDEGY-SV 122  
 QY 126 RNAIEYQVGGGINTDTLIFTDGLCDVFFVPNADOGCELVWKKSHYKHPDYCTFVENV 185  
 DB 123 KNIQMTTENNTKFNDDTVVFTDQGCVDLYIPYKEDGYELWVRSEYLNTPCQCFIDL 182  
 QY 186 FCAKDRKTYDIFNEECV 202  
 DB 183 -VALGRITYNISTPNCV 198

RESULT 4  
 AAY18085  
 ID AAY18085 standard; Protein; 203 AA.  
 XX  
 AC AAY18085;  
 XX  
 DT 06-AUG-1999 (first entry)

XX Histamine binding protein; serotonin binding compound; inflammation;  
 DE  
 KW gastric acid secretion; allergy; type I hypersensitivity reaction;  
 KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
 KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
 KW respiratory disease; coronary heart disease; cellular growth regulator;  
 KW tissue repair; blood-sucking ectoparasite; therapy.  
 XX  
 OS Boophilus microplus.  
 OS Ixodes hexagonus.  
 OS Synthetic.  
 PN WO9927104-A1.  
 XX  
 PD 03-JUN-1999.  
 XX  
 XX 26-NOV-1998; 98WO-GB03530.  
 XX  
 PR 26-JUN-1998; 98GB-0013917.  
 PR 26-NOV-1997; 97GB-0025046.  
 XX  
 PA (OXFO-) OXFORD VACS LTD.  
 XX  
 PI Nuttall PA, Paesen GC;  
 DR WPI: 1999-357841/30.  
 DR N-PSDB; AAX76970.  
 XX  
 PT Histamine and serotonin binding compounds useful for the treatment  
 of allergies  
 XX  
 PS Claim 13; Fig 7; 84pp; English.  
 CC This sequence is an example of a histamine or serotonin binding  
 compound (A), of the invention. cDNA encoding this sequence was isolated  
 from a mixed Boophilus microplus/Ixodes hexagonus cDNA expression  
 CC library. The compounds are useful for regulating the action of histamine  
 and serotonin (in e.g. inflammation and gastric acid secretion), the  
 CC detection, quantification and removal of histamine or serotonin (in  
 CC animals, plants, cell cultures, food materials, or humans) and in the  
 CC treatment of various diseases and allergies (e.g. type I hypersensitivity  
 CC reactions, urticaria, asthma, allergic rhinitis (hay fever), atopic  
 CC dermatitis, insect bites and food and drug allergies, abnormal blood  
 CC pressure, migraine, psychological disorders, respiratory disease, and  
 CC coronary heart disease). Histamine may also be used to regulate cellular  
 CC growth and tissue repair. The molecules may also be used as components of  
 CC vaccines directed against blood-sucking ectoparasites.  
 XX  
 SQ Sequence 203 AA:  
 Query Match 31.4%; Score 353; DB 20; Length 203;  
 Best Local Similarity 36.3%; Pred. No. 5.1e-27;  
 Matches 74; Conservative 41; Mismatches 81; Indels 8; Gaps 6;

QY 1 MKMOVLLITVSAALATQAEETTSKAG--ENRPLMAHEBLCKYODAMKSIDOGSVTVY 58  
 DB 1 MKALLIAGYLAALVAAQASPSRPNRLKNTW-HSKELKNODAMKSIDONVSTYY 59  
 QY 59 LAKTYENDTSMGQFKLOVQEIERKEEDYTVSVTFERNASSPIKYNTVETVKAVF 118  
 DB 60 FLRSTYNNDS-VWGNFTCLSTVTSKYSTFTVEYNTTYKNQSQ--QWVSFTEVTAQ 116  
 QY 119 QYGYKRNINAEYQVGGGINTDTLIFTDGLCDVFFVPNADOGCELVWKKSHYKHPDY 178  
 DB 117 EEGY-DVKNIQMTTENNTKFNDDTVVFTDQGCVDLYIPYKENGELWVRSDYLNTPC 175  
 QY 179 CTFVENVFCADKRTYDIFNEECV 202  
 DB 176 CQFIFDL-VALGRITYNISTPDCV 198

```

RESULT 5
AAW37447
ID AAW37447 standard; Protein: 190 AA.
XX
AC AAW37447;
XX
DT 08-JUN-1998 (first entry)
XX
DE Tick vasoactive amine binding protein 2 FS-HBP2.
XX
KW Female-specific vasoactive amine binding protein 1; FS-HCPI;
KW histamine; serotonin; assay; antihistamine; anti-inflammatory;
KW insect bite; snake bite; scorpion bite; dermatitis; vaccine;
KW transgenic animal; tick.
XX
OS Rhipicephalus appendiculatus.
XX
FH Key Location/Qualifiers
FT Peptide 1..19
FT /label= Sig_peptide
XX
PN W09744451-A2.
XX
PD 27-NOV-1997.
XX
PE 19-MAY-1997; 97WO-GB01372.
XX
PR 18-APR-1997; 97GB-0007844.
PR 18-MAY-1996; 96GB-0010484.
XX
PA (OXFO-) OXFORD VACS LTD.
XX
PI Nuttall PA, Paesen GC;
XX
DR WPI: 1998-018506/02.
DR N-PSDB; AAV00228.
XX
PT New vasoactive amine binding proteins and related nucleic acid,
PT vectors - transformed cells and transgenic animals, used for
PT assaying or removing histamine and as antihistamine or
PT anti-inflammatory agents
XX
PS Example 2; Fig 2; 44pp; English.
XX
CC This protein comprises tick Rhipicephalus appendiculatus (Ra) novel
CC female-specific histamine binding protein 2 (FS-HBP2). Its amino
CC acid sequence was deduced from a cDNA clone (see AAV00228) obtained
CC from a salivary gland cDNA library. FS-HSP1 (see AAW37446) and
CC male-specific HSP1 (see AAW37448) and a related protein, D.RET6 (see
CC AAW37449) from Dermacentor reticulatus, were also identified. These
CC novel vasoactive amine binding proteins (VABPs) can be expressed
CC in host cells using e.g. a baculovirus expression system. They can
CC be used: (i) to assay histamine (or other VA such as serotonin) in
CC body fluids or cell culture supernatants, e.g. to monitor the
CC effect of allergens; (ii) for binding VA, e.g. to remove histamine
CC from blood, food, cell cultures etc.; (iii) as an antihistamine or
CC anti-inflammatory agents, e.g. for treating insect, snake or
CC scorpion bites or dermatitis, or as a carrier for slow release of
CC histamine-related compounds; (iv) in vaccines to protect against
CC metazoan parasites, especially in animals; (v) as reagents for
CC studying inflammation, involvement of VA in ulcer formation or the
CC immune response etc. VABPs provide a more sensitive assay for
CC histamine than low-affinity antibodies currently used. They may
CC also be more effective and safer than conventional antihistamines.
XX
SQ Sequence 190 AA;
XX
Query Match 25.8%; Score 290.5; DB 19; Length 190;
Best Local Similarity 36.0%; Pred. No. 8.1e-21;
Matches 72; Conservative 32; Mismatches 79; Indels 17; Gaps 9;
XX
7 LLITFVSALATGTAETTSAGAGENPLMAHELLGKIODANKSIDGCVTVYLAKTYEN 66
II: : : II I I I I I I I I I I I I I I I I I I I I I I I I I

```

```

Db 3 LLITSLALVLA-----LSQVKGNQPDWADAEANGAHODANKSLKADVENYVYWKATYKN 57
Qy 67 DTGSMGSOFCLOQVLEERKEEDTVTSVTF-RNASSPIKYYVTEVTAVQYCKNI 125
Db 58 DP-VWGNDFTCVGMADVDEKSIQAEFLFMNADPTNQF--ATEKVTAVKMYG-NR 113
Qy 126 RNAIEYQVGGILNTDPLTFDGLCDVFVPNAD---QGCETLWKKSHKHVPDYCTFV 182
Db 114 ENAFRYETEDGQVFTDVIAYS-D-NCQVITYPGIDGNEBEYELMT--TDYDNIPANCLNR 170
Qy 183 FNFECARDKRYTDFNEECV 202
Db 171 FNEY-AVGRETRDVFTSACL 189

RESULT 6
AAV18079
ID AAV18079 standard; Protein: 190 AA.
XX
AC AAV18079;
XX
DT 06-AUG-1999 (first entry)
XX
DE Histamine binding protein FS-HBP2.
XX
KW Histamine binding protein; serotonin binding compound; inflammation;
KW gastric acid secretion; allergy; type I hypersensitivity reaction;
KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;
KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;
KW respiratory disease; coronary heart disease; cellular growth regulator;
KW tissue repair; blood-sucking ectoparasite; therapy.
XX
OS Rhipicephalus appendiculatus.
XX
PN W09927104-A1.
XX
PD 03-JUN-1999.
XX
PE 26-NOV-1998; 98WO-GB03530.
XX
PR 26-JUN-1998; 98GB-0013917.
PR 26-NOV-1997; 97GB-0025046.
XX
PA (OXFO-) OXFORD VACS LTD.
XX
PI Nuttall PA, Paesen GC;
XX
DR WPI: 1999-357841/30.
DR N-PSDB; AAX76965.
XX
PT Histamine and serotonin binding compounds useful for the treatment
PT of allergies
XX
PS Claim 1; Fig 2; 84pp; English.
XX
CC This sequence is an example of a histamine or serotonin binding
CC compound (A), of the invention. The compounds are useful for regulating
CC the action of histamine and serotonin (in e.g. inflammation and gastric
CC acid secretion), the detection, quantification and removal of histamine
CC or serotonin (in animals, plants, cell cultures, food materials, or
CC humans) and in the treatment of various diseases and allergies
CC (e.g. type I hypersensitivity reactions, urticaria, asthma, allergic
CC rhinitis (hay fever), atopic dermatitis, insect bites and food and drug
CC allergies, abnormal blood pressure, migraine, psychological disorders,
CC respiratory disease, and coronary heart disease). Histamine may also be
CC used to regulate cellular growth and tissue repair. The molecules may
CC also be used as components of vaccines directed against blood-sucking
CC ectoparasites.
XX
SQ Sequence 190 AA;
XX
Query Match 25.8%; Score 290.5; DB 20; Length 190;
Best Local Similarity 36.0%; Pred. No. 8.1e-21;

```





Oy		120	YGYNNIMNALEYOVGGGLNT--DTLFTDSELDVFPV---NADOG-CELMWKKSHK	173
Dd		107	Y---TTBNGIKYEKGTRTQTFEEVFVFSYDKNCVDLVPKRSGSDGDELMWSBKID	163
Oy		174	HVPDYCFEVENVCARDKT	193
Dd		164	KIPCCCKFTMAVFAOODEKT	183
RESULT 11				
ID	AAI18080			
ID	AAI18080 standard; Protein: 200 AA.			
XX	AAI18080:			
AC				
XX				
DT	06-AUG-1999 (first entry)			
XX				
DE	Histamine binding protein MS-HBP1.			
XX				
KM	Histamine binding protein; serotonin binding compound; inflammation;			
KM	gastric acid secretion; allergy; type I hypersensitivity reaction;			
KM	asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;			
KM	drug allergy; abnormal blood pressure; psychological disorder; vaccine;			
KM	respiratory disease; coronary heart disease; cellular growth regulator;			
KM	tissue repair; blood-sucking ectoparasite; therapy.			
OS	Rhipicephalus appendiculatus.			
XX				
PN	WO927104-A1.			
XX				
PD	03-JUN-1999.			
XX				
PF	26-NOV-1998; 98WO-GBO3530.			
XX				
PR	26-JUN-1998; 98GB-0013917.			
XX	26-NOV-1997; 97GB-0025046.			
PA	(OXFO-) OXFORD VACS LTD.			
XX				
PI	Nuttall PA, Paesen GC;			
XX				
DR	WPI: 1999-357841/30.			
XX	N-PSDB: MAX76966.			
PT	Histamine and serotonin binding compounds useful for the treatment of allergies			
PS	Claim 1; Fig 3; 84pp; English.			
XX				
CC	This sequence is an example of a histamine or serotonin binding compound (A), of the invention. The compounds are useful for regulating the action of histamine and serotonin (in e.g. inflammation and gastric acid secretion), the detection, quantification and removal of histamine or serotonin (in animals, plants, cell cultures, food materials, or humans) and in the treatment of various diseases and allergies (e.g. type I hypersensitivity reactions, urticaria, asthma, allergic rhinitis (hay fever), atopic dermatitis, insect bites and food and drug allergies, abnormal blood pressure, migraine, psychological disorders, respiratory disease, and coronary heart disease). Histamine may also be used to regulate cellular growth and tissue repair. The molecules may also be used as components of vaccines directed against blood-sucking ectoparasites.			
XX				
SQ	Sequence 200 AA:			
	Query Match 23.6%; Score 265; DB 20; Length 200;			
	Best Local Similarity 34.5%; Pred. No. 3; le-18;			
Matches	69; Conservative 34; Mismatches 71; Indels 26; Gaps 10;			
Oy	3 MOVVVLLTFVSALATQAETTSAKAGENPLAHNEELGKYODAWKSIDGVSVTVYLAKT	62		
Dd	1 MKVILLV--LCGAALCQNADA-----NPWTANENKSLSYDDAWKSLDOODKKRYLNOA	51		

QY	63	IYENITGSMGSPFKLQV--QNIERKEDYITVSTFTFNAS-SPIKYNNMTEYKAVQ	113
Db	52	TQTTD-GVWGEEFTCVSYAEKIGRKKLMATL---LYKNKHLTLDKSHETITWKAYD	106
QY	120	YGYKNRIRNAIEQVGGGLNIT--DPLIFTEGDCQVFPV---NADQG-CELMVKKSHYK	173
Db	107	Y---TTENGIKETIGTGTROTREDYFVFSYKNCVDIIFPKRGSDGDIELMVSBDKID	163
QY	174	HVPDYCTEFVFNVCARDKRT	193
Db	164	KIPDCCKFTMAYFAOQOEXT	183
RESULT 12			
AAW37446			
ID	AAW37446	standard; Protein; 190 AA.	
XX	AAW37446;		
AC			
XX	08-JUN-1998	(first entry)	
XX			
DE	Tick vasoactive amine binding protein, 1 FS-HBP1.		
XX			
KW	Female-specific vasoactive amine binding protein 1; FS-HCP1;		
KW	histamine; serotonin; assay; antihistamine; anti-inflammatory;		
KW	insect bite; snake bite; scorpion bite; dermatitis; vaccine;		
KW	transgenic animal; tick.		
XX			
OS	Rhipicephalus appendiculatus.		
XX			
FH	Key	Location/Qualifiers	
FT	Peptide	1..18	
FT		/label= Sig_peptide	
XX			
PN	WO9744451-A2.		
XX			
PD	27-NOV-1997.		
XX			
PF	19-MAY-1997;	97WO-GB01372.	
XX			
PR	18-APR-1997;	97GB-0007844.	
PR	18-MAY-1996;	96GB-0010484.	
XX			
PA	(OXFO-) OXFORD VACS LTD.		
XX			
PI	Nuttall PA, Paesen GC;		
XX			
DR	WPI: 1998-018506/02.		
DR	N-PSDB: AAW00227.		
XX			
PT	New vasoactive amine binding proteins and related nucleic acid,		
PT	vectors - transformed cells and transgenic animals, used for		
PT	assaying or removing histamine and as antihistamine or		
PT	anti-inflammatory agents		
XX			
PS	Example 2; Fig 1; 44pp; English.		
XX			
CC	This protein comprises tick Rhipicephalus appendiculatus (Ra) novel		
CC	female-specific histamine binding protein 1 (FS-HBP1). Its amino		
CC	acid sequence was deduced from a cDNA clone (see AAW00227) obtained		
CC	from a salivary gland cDNA library. FS-HSP2 and male-specific		
CC	HSP1 (see AAW37447-48) and a related protein, D-RET6 (see AAW37449)		
CC	from Dermacentor reticulatus, were also identified. These novel		
CC	vasoactive amine binding proteins (VABPs) can be expressed in		
CC	host cells using e.g. a baculovirus expression system. They can		
CC	be used: (i) to assay histamine (or other VA such as serotonin) in		
CC	body fluids or cell culture supernatants, e.g. to monitor the		
CC	effect of allergens; (ii) for binding VA, e.g. to remove histamine		
CC	from blood, food, cell cultures etc.; (iii) as an antihistamine or		
CC	anti-inflammatory agents, e.g. for treating insect, snake or		
CC	scorpion bites or dermatitis, or as a carrier for slow release of		
CC	histamine-related compounds; (iv) in vaccines to protect against		



Db 60 LNATI-----LYNNKHLTDLKESHEITTYMKAYDI-----TTENGIGIKETGOTGTOTIFEDYFV 112

Qy 145 FIDGELCDYFYPP---MADQG-CELMVKKSHKKHYPDCTCFVNFPCAIDKRT 193  
113 FSYKKKCDYIFPFKENGSDGDEYELMAVSDKIDTKPDCCKTMMAYFAADQCKT 165

Search completed: August 4, 2003, 17:14:39  
Job time : 84 secs

88 EDYTVTSVFERNAS-SPIKYYNVTETVKAVFQYGYKNIRNAIEYQVGGGLNIT--DTLI 144

Db 60 LNATI-----LYKNKHLTDLKESHETITVWKAYDY--TTENGIKYETOGTRTQTFEDV FV 112

```

Qy      145  FTDEGLCEDEVFP--NADOG-CELMVKKSHKHVPDYCTFVENVFCANDRKT 1933
        |:| : |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db      113  FSDYKNCVDVIFVPKERGSEGDYELMVSEDIKIPCCCKFTMAVFAQOQEXT 1655

```

Search completed: August 4, 2003, 17:14:39  
Job time : 84 secs

THIS PAGE BLANK (USPTO)



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: August 1, 2003, 12:15:04 ; Search time 52.3467 Seconds  
(without alignments)  
897.203 Million cell updates/sec

Title: US-10-087-195-8  
Perfect score: 989  
Sequence: 1 NPTWANEAKIGSYQDAMKSL.....EKTVRNYTDSCKPAPAQN 182

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

SPTREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_protist:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriophage:\*
- 17: sp\_archaeal:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	251.5	25.4	209	5	Q8WSK7 dermatocentor
2	113.5	11.5	330	3	Q05583 saccharomyc
3	108.5	11.0	311	5	Q8MVC2 ixodes scap
4	95.5	9.7	210	5	Q8MVC2 ixodes scap
5	92	9.3	220	5	Q8MVC2 ixodes scap
6	92	9.3	306	5	Q8MVC2 ixodes scap
7	92	9.3	581	4	Q8MVC2 ixodes scap
8	92	9.3	956	5	Q8MVC2 ixodes scap
9	92	9.3	169	2	Q05273 bacillus su
10	87	8.8	483	16	Q928X7 listeria su
11	87	8.8	483	16	Q928X7 listeria su
12	86.5	8.7	604	11	Q8K3S1 mus musculu
13	86.5	8.7	3467	5	Q8K3S1 mus musculu
14	86	8.7	391	16	Q8FQ68 corynebacte
15	85.5	8.6	428	16	Q8F3F8 leptospira
16	85.5	8.6	1543	5	Q8MGY7 drosophila

17	85.5	8.6	1970	5	Q9VCH4	Ogyc4 drosophila
18	85	8.6	514	10	Q8L6A8	Q8L6A8 theobroma c
19	85	8.6	662	5	Q61096	Q61096 trypanosoma
20	85	8.6	1959	16	Q8YANI	Q8YANI listeria mo
21	84	8.5	582	10	Q43079	Q43079 plasm saliv
22	84	8.5	1073	5	Q8LJWS	Q8LJWS plasmodium
23	83.5	8.4	778	2	Q9EXN1	Q9EXN1 enterococcu
24	83.5	8.4	1078	12	Q90193	Q90193 agaricus bl
25	83	8.4	1124	5	Q8LIS3	Q8LIS3 plasmodium
26	83	8.4	3896	12	P87314	P87314 border dise
27	82.5	8.3	306	9	Q64070	Q64070 bacterioph
28	82.5	8.3	306	16	Q31951	Q31951 bacillus su
29	82.5	8.3	498	16	Q8F3P9	Q8F3P9 leptospira
30	82.5	8.3	1802	5	Q96170	Q96170 plasmodium
31	82	8.3	392	16	Q8EAC0	Q8EAC0 streptococc
32	82	8.3	392	16	Q8DIV2	Q8DIV2 streptococc
33	82	8.3	577	16	Q8E147	Q8E147 shewanella
34	82	8.3	617	16	Q8X116	Q8X116 clostridium
35	81.5	8.2	196	16	Q8R7C5	Q8R7C5 thermomater
36	81.5	8.2	388	16	Q8R1M4	Q8R1M4 fusobacteri
37	81.5	8.2	501	16	Q8PQC5	Q8PQC5 xanthomonas
38	81.5	8.2	817	16	Q93214	Q93214 staphylococ
39	81	8.2	223	5	Q9GP33	Q9GP33 echinococcu
40	81	8.2	581	6	Q95LY3	Q95LY3 macaca fasc
41	81	8.2	912	2	Q06915	Q06915 lactobacill
42	81	8.2	953	16	Q8EYF8	Q8EYF8 mycoplasma
43	80.5	8.1	433	17	Q97AY2	Q97AY2 thermoplasma
44	80.5	8.1	641	5	Q05508	Q05508 trypanosoma
45	80.5	8.1	721	16	Q8Y763	Q8Y763 listeria mo

## ALIGNMENTS

## RESULT 1

Q8WSK7 PRELIMINARY: PRT; 209 AA.

Q8WSK7: 01-MAR-2002 (TREMBLrel. 20, Created)

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Serotonin and histamine binding protein.

OS Dermacentor reticulatus.

OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;

OC Parasitiformes; Ixodidae; Ixodidae; Dermacentor.

OX NCBI\_TaxID=57047;

RN (1)

RP SEQUENCE FROM N.A.

RA Sangamadech S., Paesen G.C., Nuttall P.A.;

RT "A high affinity serotonin- and histamine-binding lipocalin secreted by blood-feeding ticks."

RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF217101; AAL5664.1; -

DR InterPro: IPR002970; His\_binding.

DR Pfam: PF02098; His\_binding; 1.

SO SEQUENCE 209 AA; 23889 MW; E5EB8BE710859591 CRC64;

Query Match 25.4%; Score 251.5; DB 5; Length 209;

Best Local Similarity 35.3%; Pred. No. 2,1e-14;

Matches 61; Conservative 29; Mismatches 66; Indels 17; Gaps 8;

QY 1 NPTWANEAKIGSYQDAMKSLQDDDKRRYLAQATPTD-GVWGGEFTCVSYAEKIKKK 59

DB 30 NPLWAHEELTKYQDAMKSLDQGVSVTVLVLAFTYENDTGSWSQKCLQV--QETEKKE 87

QY 60 LNATL---LYKNKHLTDLKSHEHETIYKAKADV---TTENGIRKETGTGTQTPQEDYFV 112

DB 88 EDYITVSVFTFRNMS-SPIKYVNTVEYKAVQYQYKIRNRIAEIQVGGGLNIT--DPLI 144

QY 113 FSDYKNCVIFVPERGSDGDEYELWSEDKIDKIPDCCKFTMAVFAQQQEXT 165

DB 145 FTDGELCDVFPYV---NADQG-CELMVKKSHYKHVPDYCTFVFNFCCKDKRRT 193

RESULT 2  
ID 005583 PRELIMINARY; PRT: 330 AA.  
AC 005583:  
DB 01-NOV-1996 (TREMblrel. 01, Created)  
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
DE 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
DE Similar to S. CEREVISIAE SEC13P.  
GN YDR267C OR D9954.11.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C;  
RA Le T.;  
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C;  
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,  
RA Favell A., Fulton L., Gattung S., Greco T., Kirsten J., Kucada T.,  
RA Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D.,  
RA Johnston L., Langston T., Lacroille P., Le T., Mardis E., Menezes S.,  
RA Miller N., Nhan M., Pauley A., Peluso D., Rifkin L., Riles L.,  
RA Talch A., Trevasakis E., Vignati D., Wilcox L., Wohlman P., Vaudin M.,  
RA Wilson R., Waterston R.;  
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C;  
RA Waterston R.;  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C;  
RA Jia Y., Cherry J.M.;  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U51030; AAB64456.1;  
DR SGD: S0002675; YDR267C.  
DR InterPro: IPR001680; WD40.  
DR Pfam: PF00400; WD40; 7.  
DR PRINTS: PR00320; GPROTEINRPT.  
DR ProDom: PD000018; WD40; 1.  
DR SMART: SM00320; WD40; 7.  
DR PROSITE: PS00678; WD\_REPEATS\_1; 1.  
DR PROSITE: PS50082; WD\_REPEATS\_2; 5.  
DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
KW Repeat; WD repeat.  
SQ SEQUENCE 330 AA; 37274 MW; B44CCAA3125FD666 CRC64;  
Query Match 11.5%; Score 113.5; DB 3; Length 330;  
Best Local Similarity 27.5%; Pred. No. 0.051;  
Matches 38; Conservative 17; Mismatches 62; Indels 21; Gaps 5;  
QY 16 AMKSLQDQDKRRYLAQATQTTDGVW-----GEFTCVSYAEKIGKKLNATILYKN 68  
DB 114 AM-----SNDGYLLATCSR-DKSVWIMETDESGEEYECISVLQEH--SQDVKHVIMHPS 164  
QY 69 KHLTLKESHETTIVKAYITTE-----NGIYETOGTITQTFEDVYFVSQKNCVIF 123  
DB 165 EALLASSSYDDTVIRIMWDYDDDWECVAVLHGEGTWSDFTEGVFRLCSGSDSTVR 224  
QY 124 VPKRGSDEGDYELWSE 141  
DB 225 YKYMKGDEDDDDQEWCE 242

AC 08MVC2;  
DT 01-OCT-2002 (TREMblrel. 22, Created)  
DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)  
DE 01-OCT-2002 (TREMblrel. 22, Last annotation update)  
DE Putative secreted histamine binding protein.  
OS Ixodes scapularis (Black-legged tick) (Deer tick).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Parasitiformes; Ixodidae; Ixodidae; Ixodes.  
OX NCBI\_TaxID=6945;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Rhode Island; TISSUE-Salivary gland;  
RA Valenzuela J.G., Francischetti I.M., Pham V.M., Garfield M.,  
RA Mather T.N., Ribeiro J.M.C.;  
RT "Exploring the Salivome of the Tick Vector of Lyme Disease, Ixodes  
scapularis".  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF483742; AAM93640.1;  
SQ SEQUENCE 311 AA; 36581 MW; D7E529FEFEACFA9A CRC64;

Query Match 11.0%; Score 108.5; DB 5; Length 311;  
Best Local Similarity 24.3%; Pred. No. 0.13;  
Matches 46; Conservative 25; Mismatches 65; Indels 53; Gaps 9;  
QY 1 NPTVANEAKLSYODANKSLQDQDKRRYLAQATQTTDGVW-----EFTCVSYAEKI 55  
DB 32 NPD-LNKRKDLGAMQDMARTIKFTANHSYLLYSFG-----WGTRHEYEDVRCLOVHSSDL 85  
QY 56 G-KKKLNATILYKNKHLTLKESHETTIVKAYITTE-----TTENGIKY- 97  
DB 86 NYTLKSANYSKWNRKNSKRNSSYVOAAKORYSIENIMHLPQREYTSNGCYN 145  
QY 98 -----ETQGT-----TQTFEDVYFVSQKNCVIFPKERSDEG--DYEL 137  
DB 146 LNFENFLESQSGCRHHQECMOKRWTXYSEKYLFS-----TPLCVNLSLDDDEYSCSEF 201  
QY 138 WSEEDKTDK 146  
DB 202 WLSEEDLTK 210

RESULT 4  
ID 08MVC2 PRELIMINARY; PRT: 210 AA.  
AC 08MVC2:  
DT 01-OCT-2002 (TREMblrel. 22, Created)  
DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)  
DE 01-OCT-2002 (TREMblrel. 22, Last annotation update)  
DE Putative 22.5 kDa secreted protein.  
OS Ixodes scapularis (Black-legged tick) (Deer tick).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Parasitiformes; Ixodidae; Ixodidae; Ixodes.  
OX NCBI\_TaxID=6945;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Rhode Island; TISSUE-Salivary gland;  
RA Valenzuela J.G., Francischetti I.M., Pham V.M., Garfield M.,  
RA Mather T.N., Ribeiro J.M.C.;  
RT "Exploring the Salivome of the Tick Vector of Lyme Disease, Ixodes  
scapularis".  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF483742; AAM93640.1;  
SQ SEQUENCE 210 AA; 24689 MW; C5B75614649A9BED CRC64;  
Query Match 9.7%; Score 95.5; DB 5; Length 210;  
Best Local Similarity 22.4%; Pred. No. 1.1;  
Matches 38; Conservative 30; Mismatches 59; Indels 43; Gaps 9;  
QY 7 EARLGSYODANKSLQDQDKRRYLAQATQTTDGVWGEFTCVS-----V 50  
DB 29 DEKTYMOYQDIORLNNDPRESWLYRYRRETQ--GSEHICSAKYSQNGSDYEFV 85  
QY 51 TAEKIGKKLN---ATILYKNKHLTL-----KESHETTIVKAYDYTTENGIKYETQGT 103

Db 86 GERYLGTKEQNRKRTVTLATATPYKTEMHATQRONNAMVSOAKD--ADGKGXYQ----- 138  
QY 104 TTFEDVVFVSDYKNCDFVFKRSGDEGYELWVSEDKIDK-IPDCK 152  
139 -----LIYSDYKNCIDLRLVRENSGH--DCELYLHRSKALDDGVPRECE 179

RESULT 5  
ID Q8MVB8 PRELIMINARY: PRT; 220 AA.  
AC Q8MVB8; 01-OCT-2002 (Tremblrel. 22, Created)  
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
DE 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
OC Putative secreted protein.  
OS Ixodes scapularis (Black-legged tick) (Deer tick).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Parasitiformes; Ixodida; Ixodidae; Ixodes.  
OX NCBI\_TaxID=6945;

SEQUENCE FROM N.A.  
RC STRAIN-Rhode Island; TISSUE-Salivary gland;  
RA Valenzuela J.G., Francischelli I.M., Pham V.M., Garfield M.,  
RT Mather T.N., Ribeiro J.M.C.;  
RT "Exploring the salivome of the tick vector of Lyme Disease, Ixodes  
scapularis";  
RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AF483722; AAM93644.1; -  
SQ SEQUENCE 220 AA; 25375 MW; 61ADD7258AA28DDE CRC64;

Query Match  
Best Local Similarity 23.7%; Pred. No. 2.5; Length 220;  
Matches 47; Conservative 29; Mismatches 70; Indels 52; Gaps 12;

QY 6 NEAKLSYODAMKSLQODQNKRRYLAQATQTDGVWGEFVCVSYTA----- 52  
Db 46 NSAKL-----WL-----KRRKFLYVRSKANDTERGCGACQVOKKAFIELPKRRFSWN 94  
QY 53 EIKGKKLNLATILYKKNHLTDLKESHETTVKAVDYTTENGICKETG---GTROT--- 106  
Db 95 TPVTLKLVFSSVGRNNGTGHNDGFTALITTKSYE---GLVYDDMFSGIELYTHYI 149  
QY 107 FEDV-FVFSDDYKNCDFVFKRSGDEGYELWVSEDK-IDKIPDCKFTMAYFAQDE 163  
Db 150 YVDYELFTDYETCTTMRP---SDD-VYVMWMDRRNITIDINPQC---ETAYHGPVNE 201  
QY 164 -----KTVRNVYTDSSC 175  
Db 202 YGCTVAVKPKRYVYDEMIC 219

RESULT 6  
ID Q8MVC3 PRELIMINARY: PRT; 306 AA.  
AC Q8MVC3; 01-OCT-2002 (Tremblrel. 22, Created)  
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
DE 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
OC Putative secreted histamine binding protein.  
OS Ixodes scapularis (Black-legged tick) (Deer tick).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Parasitiformes; Ixodida; Ixodidae; Ixodes.  
OX NCBI\_TaxID=6945;

SEQUENCE FROM N.A.  
RC STRAIN-Rhode Island; TISSUE-Salivary gland;  
RA Valenzuela J.G., Francischelli I.M., Pham V.M., Garfield M.,  
RT Mather T.N., Ribeiro J.M.C.;  
RT "Exploring the salivome of the tick vector of Lyme Disease, Ixodes  
scapularis";  
RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AF483717; AAM93639.1; -

SQ SEQUENCE 306 AA; 34232 MW; B2DB09181CC4395D CRC64;  
Query Match  
Best Local Similarity 9.3%; Score 92; DB 5; Length 306;  
Matches 42; Conservative 25; Mismatches 61; Indels 26; Gaps 10;

QY 1 NPTWANEAKLSYODAMKSLQODQNKRRYLAQATQTDGVWGEFVCVSYTAEKICK--K 58  
Db 30 NPL-LNEPSLGPQSAKAIKNSKDDPFLMFRSRNH-----PNTCVVVTASLHNETLK 84  
QY 59 KLNATILYKKNHLTDLKE--HETTVWKAYDYTTENGICKYETOGT---RTQFEDVVF 112  
Db 85 IYVNFRTTYNE--TTGKNNTLEYQVRAALNQTYKLENTVIRAGLKGTPSDKPPILGSMNI 142  
QY 113 -FSDYKNC-----DVIFPKER-GSDEGYELW 136  
Db 143 EYGDY-SCNTSSKRPYLDMLKAKDAVGSBAGSAE 175

RESULT 7  
ID Q8NA03 PRELIMINARY: PRT; 581 AA.  
AC Q8NA03; 01-OCT-2002 (Tremblrel. 22, Created)  
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
DE 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
OC Hypothetical protein FLJ35989.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;

SEQUENCE FROM N.A.  
RC TISSUE-Testis;  
RA Suzuki O., Sasaki N., Aotsuka S., Shoji T., Ichihara T., Shiohata N.,  
RA Matsumoto K., Hirano M., Sano S., Nomura R., Yoshikawa Y.,  
RA Matsumura Y., Moriya S., Chiba E., Nomura R., Onogawa S.,  
RA Kaeriyama S., Satoh N., Matsunawa H., Takahashi E., Kataoka R.,  
RA Kuga N., Kuroda A., Satoh I., Kamata K., Takami S., Terashima Y.,  
RA Watanabe M., Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A.,  
RA Ishii S., Yamamoto J., Isono Y., Kawaki-Hio Y., Saito K., Nishikawa T.,  
RA Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,  
RA Kikuchi H., Kanda K., Magatsuma M., Murakawa K., Kanehori K.,  
RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,  
RA Sugano S., Nagahara K., Masuno Y., Nagai A., Isogai T.;  
RT "NEO human cDNA sequencing project";  
RL Submitted (JUL-2002) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AK093308; BAC04128.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 581 AA; 66121 MW; 979A8B4DE217C74 CRC64;

Query Match  
Best Local Similarity 9.3%; Score 92; DB 4; Length 581;  
Matches 40; Conservative 25; Mismatches 66; Indels 28; Gaps 7;

QY 4 WANEAKLSYODAMKSLQODQNKRRYLAQATQTDGVWGEFVCVSYTAEKICKRL 60  
Db 151 W-EETKSAYSEAWOSKEEMTKRFLSLAASEEIVGSHBEDTFSSVFTQIAPPEY 209  
QY 61 NATIILYKKNHL-DKESHETTVKAYDYTTENGICKYETOGTROT---EDVVFSDYK 117  
Db 210 ENQMKRLNDFCDVBRNLSLKGKPPRSNF---KILRKHNDGFIKRIELAKESR 266  
QY 118 NCDVIFPKER-----GSDEGYELW 139  
Db 267 N-PVWVDRKKRLVELLKDLDKDSGLSSSEGSQGW 304

RESULT 8  
ID Q8IB10 PRELIMINARY: PRT; 956 AA.  
AC Q8IB10; 01-MAR-2003 (Tremblrel. 23, Created)

DT 01-MAR-2003 (TReMBLrel. 23, last sequence update)  
DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)  
DE Hypothetical protein.  
GN PF07\_0114.  
OS Plasmidium falcatiparum (isolate 3D7).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.  
OX NCBI\_TaxID=36329;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Seeger K., Murphy L., Harris D., Berriman M., Pain A., Hall N.,  
RA Quail M., Barrrell B.;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL844506; CAD51020.1; -  
KM Hypothetical protein.  
SQ SEQUENCE 956 AA; 114829 MW; BF81CA938F55B0A CRC64;  
Query Match 9.3%; Score 92; DB 5; Length 956;  
Best Local Similarity 28.2%; Pred. No. 15;  
Matches 33; Conservative 13; Mismatches 35; Indels 36; Gaps 6;  
QY 58 KKLNATLYNKKHLTDI--KESHTTVMKAYDYTTENGIKYGTGTGTGFEDVFEVSD 115  
DB 24 KKYNEFEFFYKNTKLKDKSKRENNHFDIYKLY-----EEKNYITD 63  
QY 116 YKNDVTFVFKERGSDEGVLEWVSEDKIDKIPDCKFTMAYFAQOQEKTVRNYTD 172  
DB 64 Y--CNLLV-----DEIDNSFKIETDN-KSHNICK---YPLQNFDEWNTNNTYD 106  
RESULT 9  
005273 PRELIMINARY; PRT; 169 AA.  
AC 005273;  
DT 01-JUL-1997 (TReMBLrel. 04, Created)  
DT 01-JUL-1997 (TReMBLrel. 04, last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, last annotation update)  
DE Hypothetical 19.8 kDa protein.  
GN YUCC.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA SPRAIN-168;  
RA Oudega B., Koningsstein G., Duisterhoef A.;  
RA Bacillus subtilis genome project, DNA sequence from yucca to yuch.;  
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: Z93940; CAB07966.1; -  
KM Hypothetical protein.  
SQ SEQUENCE 169 AA; 19833 MW; 211319D174AF435F CRC64;  
Query Match 8.8%; Score 87; DB 2; Length 169;  
Best Local Similarity 26.9%; Pred. No. 4.9;  
Matches 35; Conservative 19; Mismatches 48; Indels 28; Gaps 6;  
QY 45 FT-CVSTAEKIKKKLNA--TILYKKNKHLTDK-----ESHETITVMKAYDYTTENG 94  
DB 26 FTWCCHQLTTEELKKNLKGVEVNLITLPHITIDVEYSWIRALQCKEDIAVOFA-DYOTLTK 84  
QY 95 IKYETGTGTQPEDEFEVSDYKNCVDVIFPKERG-----SDE-----GDYEL 137  
DB 85 VKSLSMTFTREITIDVQTHSDQIKDELVSVPWETGVLYTRDELHHTIHAHEIHIGOLSY 144  
QY 138 WVEDKIDKI 147  
DB 145 WARELKLSPV 154  
RESULT 10  
092BX7 PRELIMINARY; PRT; 483 AA.  
AC 092BX7;  
DT 01-DEC-2001 (TReMBLrel. 19, Created)

DT 01-DEC-2001 (TReMBLrel. 19, last sequence update)  
DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)  
DE Two-component sensor histidine kinase.  
GN LISK OR LINI415.  
OS Listeria innocua.  
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
OX NCBI\_TaxID=1642;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN-CLIP 11262 / Serovar 6a;  
RX PubMed=11679669;  
RA Glaser P., Franjeul L., Buchrieser C., Rusniok C., Amend A.,  
RA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,  
RA Charbit A., Chehouani F., Couve E., de Daruvar A., Dehoux P.,  
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurgeat O.,  
RA Entian K.-D., Feihl H., Garcia-del Portillo F., Garrido P.,  
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kutupat G.,  
RA Madueno E., Maitounan A., Mata Vicente J., Ng E., Nedjari H.,  
RA Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,  
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,  
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;  
RL Science 294:849-852 (2001).  
DR EMBL: AL596168; CAC96646.1; -  
DR L1stList; LIN01415; -  
DR InterPro: IPR003594; ATPbind\_ATPase.  
DR InterPro: IPR004358; Bact\_sens\_pr\_C.  
DR InterPro: IPR003660; HAMp.  
DR InterPro: IPR003661; His\_kinA.  
DR InterPro: IPR005467; His\_kinase.  
DR Pfam: PF00672; HAMp; 1.  
DR Pfam: PF02518; HAMpase\_c; 1.  
DR Pfam: PF00512; HisKA; 1.  
DR PRINTS: PR00344; BCTRLSENSOR.  
DR SMART: SM00304; HAMp; 1.  
DR SMART: SM00387; HAMpase\_c; 1.  
DR SMART: SM00388; HisKA; 1.  
DR PROSITE: PS50109; HIS\_KIN; 1.  
KM Kinase; Complete proteome.  
SQ SEQUENCE 483 AA; 55501 MW; F5712DF5304C32D2 CRC64;  
Query Match 8.8%; Score 87; DB 16; Length 483;  
Best Local Similarity 24.0%; Pred. No. 18;  
Matches 37; Conservative 29; Mismatches 58; Indels 30; Gaps 7;  
QY 18 KSIQODNKKRYLQAOTGTGDVGWGEFTCVSTAERIGKKKLNATLYNKKHLTDKES 77  
DB 308 KKLVDKMLDSRADQISQTE---LQITVNAVTEGV---RRNFVMTENFTFT-LKED 359  
QY 78 HETITVMKAYDY-----TTENGIKYETGTGTQTFEDVFEVSDYKNCVDVIFPKERGS 130  
DB 360 DTDLRALQNHNLQEQILITIDNAVKVSGDTEV---DMHYVKEQKHNI----- 406  
QY 131 DEGDYELWVSEDKIDKIPDCKFTMAYFAQOQEK 164  
DB 407 DVRDYGEIGISQIEIDKIFN--RFRVYDKARSREK 438  
RESULT 11  
09RPY9 PRELIMINARY; PRT; 483 AA.  
AC 09RPY9;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, last sequence update)  
DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)  
DE LISK (Two-component sensor histidine kinase).  
GN LISK OR LMO1378.  
OS Listeria monocytogenes.  
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
OX NCBI\_TaxID=1639;  
RN [1]  
RP SEQUENCE FROM N.A.

RA STRAIN-1028 / Serovar 1/2c:  
RX MEDLINE=20011355; PubMed=10542190:  
RT "identification and disruption of *lspK*, a genetic locus encoding a  
RT two-component signal transduction system involved in stress tolerance  
RT and virulence in *Listeria monocytogenes*."  
RL J. Bacteriol. 181:6840-6843(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-EGD-e / Serovar 1/2a:  
RX MEDLINE=21537279; PubMed=11679669;  
RA Glaser P., Frangeul L., Buchrieser C., Rusnlok C., Amend A.,  
RA Baquero F., Berche P., Bloecher H., Brandt P., Chakraborty T.,  
RA Charbalt A., Chetoui F., Couve E., de Daruvar A., Deloux P.,  
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
RA Gantier K.-D., Fshli H., Garcia-del Portillo F., Garrido P.,  
RA Gauthier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kuropat G.,  
RA Madueno E., Maitouram A., Mata Vicente J., Ng E., Nedjati H.,  
RA Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,  
RA Remmel B., Rose M., Schluteter T., Simoes N., Tietz A.,  
RA Vaqueria-Boland J.-A., Voss H., Weiland J., Cossart P.,  
RT "Comparative genomics of *Listeria* species".  
RL Science 294:849-852(2001).  
CC -1- SIMILARITY: TO OTHER PROKARYOTIC SENSOR TRANSDUCTION HISTIDINE  
CC KINASES.  
CC EMBL: AF193908; AF039333.1; -  
CC EMBL: AL591978; CAC99456.1; -  
CC ListiList: LMO01378; -  
DR InterPro: IPR003594; ATPbind\_ATPase.  
DR InterPro: IPR004358; Bact\_sens\_pr\_C.  
DR InterPro: IPR003660; HAMP.  
DR InterPro: IPR003661; His\_kinae.  
DR InterPro: IPR005467; His\_kinae.  
DR Pfam: PF00672; HAMP; 1.  
DR Pfam: PF02518; HATPase\_c; 1.  
DR Pfam: PF00512; HSKA; 1.  
DR PRINTS: PR00344; BCTRLSENSOR.  
DR SMART: SM00304; HAMP; 1.  
DR SMART: SM00387; HATPase\_c; 1.  
DR SMART: SM00388; HSKA; 1.  
DR PROSITE: PS50109; HIS\_KIN; 1.  
KW kinase: Phosphorylation; Sensory transduction; Transferase;  
KW Complete proteome.  
KW SEQUENCE 483 AA; 55460 MW; B477C2C3D653DC95 CRC64;  
Query Match 8.8%; Score 87; DB 16; Length 483;  
Best Local Similarity 24.0%; Pred. No. 18;  
Matches 37; Conservative 29; Mismatches 56; Indels 30; Gaps 7  
QY 18 KSLQDQDKRRYYLAQATQTDTGDSVGEFTCVSYTAETKIGKRLNATILYKNNHLDKES 77  
Db 308 KKLVEQMDLSRAEGISQTK-----LQITVDNATVEGV---RRNEVYENTPFL-LKED 359  
QY 78 HETTYVMKAYDY-----TTENGKIKYTGCTGTOTFEEDVFSYKNCVDIYFKERGS 130  
Db 360 DTDLRALQHNHLEQILIIINDNVAKYSGGDGEV---DMHYEKQKQIHI----- 406  
QY 131 DEGDVEYLVSEDKIDKIDPCCKFTPMAYFAAOOEK 164  
Db 407 DVYDYGEGISQDEIDKIFN--RFTYVDKARSREK 438  
ID 08K3S1 PRELIMINARY; PRT; 604 AA.  
AC 08K3S1  
DT 01-OCT-2002 (TREMBlrel. 22, Created)  
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE PpxX protein.  
GN PpxX.  
OS Mus musculus (Mouse).

```

OC Eukaryota;Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=22040331; PubMed=12045143;
RA Chureau C., Prissette M., Bourdet A., Barbe V., Cattolico L.,
RT Jones L., Eggen A., Averb P., Duret L.;
RT "Comparative sequence analysis of the X-inactivation center region in
RL mouse, human and bovine.";
RL Genome Res. 12:894-908(2002).
DR EMBL; AJ421480; CAD33962.1; -
DR InterPro; IPR001254; Ser_protease-Try.
DR SMART; SM00020; Tryp_Spc. 1.
DR PROSITE; PS50240; TRYPSIN_DOM. 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 604 AA; 67709 MW; 5542092EBFEA8ACD CRC64;

Query Match 8.7%; Score 86.5; DB 11; Length 604;
Best Local Similarity 26.7%; Pred. No. 25;
Matches 27; Conservative 14; Mismatches 41; Indels 19; Gaps 4;

OY 24 ONKRYLAQAQTGDSVGMEPTCVSYAEIKGKKLNAAT-----ILYKNKLHLDLKS 77
Db 234 ONPAVITTAAYFT---WIKQFI-----KYSKLLPMTSLHCRRTLYENEYQPIRHS 283
OY 78 HETIYWKAYDY---TTENGIRKETQGTROTFFEDVEFVSF 115
Db 284 HNFFYTLASSDYLLRLILGNGVSPQTHTNVPIIFNPSD 324

RESULT 13
O8I218 PRELIMINARY; PRT; 3467 AA.
ID O8I218;
AC O8I218;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Erythrocyte membrane protein 1 (PFEMP1).
GN VAR.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255708; PubMed=12368867;
RA Hall N., Pain A., Barriman M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA Fellwell T., Goble A., Goodhead I., Gilliam R., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA Humphrey S., Jorgels K., James K.D., Johnson D., Kethonou A.,
RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
RA Line A., Maddison M., Mclean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sulston J.E., Craig A., Newbold C., Barrell B.G.;
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
RL Nature 419:527-531(2002).
DR EMBL; AL034557; CADA9096.1; -
SQ SEQUENCE 3467 AA; 398212 MW; E8FCBA37025989D2 CRC64;

Query Match 8.7%; Score 86.5; DB 5; Length 3467;
Best Local Similarity 21.7%; Pred. No. 21e+02;
Matches 46; Conservative 30; Mismatches 63; Indels 73; Gaps 12;

OY 15 DAWKSL-----QDQDNK---RYLAQAQTGTTDGV-----WGEE 44

```

DB 1375 EIMEGMLCALTHEIDEEKKIKSTYSYDOLAKTTNGTPTLEKFAERQFLRWTEWSE 1434  
OY 45 FTVSVTAERKIGKKLNATILYKNNHLDLTKESHETTVMKAY-DYTTENGIRYETQGT 103  
DB 1435 FCRERREKEVEERKCKKDHGCKNKPNT--KGNHCVSACADYEYISTKCKKQYNTQ--- 1489  
OY 104 TQTFEDVFVPSDYKNCVDYFVFKERGSDEGDEYELWVSDKIDKIPD-CCKRTMAYFAOQ 162  
DB 1490 KKKFD-----IDKKNKNGEE--YENYKDKKKAHDYLDKCKPCTCDY---M 1528  
OY 163 EKVTVN-----VYVDS-----SCKPAP 179  
DB 1529 EKVKNSEYMDKPKNTYTNSDLEKKCKECPKP 1560  
RESULT 14  
O8F068  
ID 08F068 PRELIMINARY; PRT; 391 AA.  
AC 08F068:  
DT 01-MAR-2003 (TREMblrel. 23, Created)  
DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)  
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
DE Conserved hypothetical protein.  
GN CE1265.  
OS Corynebacterium efficiens.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
OX NCBI\_TaxID=152794;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;  
RA Kawarabayasi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,  
RA Ito K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,  
RA Usuda Y., Sugimoto S.;  
RT "The entire genomic sequence of Corynebacterium efficiens YS-314.";  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AP005218; BAC18075.1;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 391 AA; 42871 MW; 07C4C6D4F41F0 CRC64;  
Query Match 8.7%; Score 86; DB 16; Length 391;  
Best Local Similarity 27.8%; Pred. No. 17;  
Matches 40; Conservative 15; Mismatches 43; Indels 46; Gaps 7;  
OY 27 RYVLAQTQTTD-----GYMGE--EFTCVSVTAEK-----IGKKLNATILYKN 68  
DB 218 RVWFSGAPETDDEHLDPTRVGGEVNSGNLVYVASKGHAVEGVPGKWFEAL-----N 273  
OY 69 KHLTDLKESHETTVMKAYVTTENGIRYETQGT-----QTFEDVFVPSDYK 117  
DB 274 RELTSDIEDIEFTLTLRAVPPDKSTVIRYGLQGTITLQNNRRLREGLASLEDIF----- 327  
OY 118 NCDVIFVFKERGS-----EGDYEL 137  
DB 328 ---ASLKRERTDVLVLEPDEEL 348  
RESULT 15  
O8F3F8  
ID 08F3F8 PRELIMINARY; PRT; 428 AA.  
AC 08F3F8:  
DT 01-MAR-2003 (TREMblrel. 23, Created)  
DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)  
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
DE Putative outer membrane protein.  
GN LA2448.  
OS Leptospira interrogans.  
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.  
OX NCBI\_TaxID=173;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar 1a1;  
RA Ren S.;

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AE011413; AAN9647.1;  
KW Complete proteome.  
SQ SEQUENCE 428 AA; 48954 MW; 0F7DD095D1BA1C7C CRC64;  
Query Match 8.6%; Score 85.5; DB 16; Length 428;  
Best Local Similarity 22.9%; Pred. No. 21;  
Matches 38; Conservative 31; Mismatches 60; Indels 37; Gaps 6;  
OY 5 ANEAKLSYODAMKSLQDDQNKRRYLAQTQTTGVNGEFTCVSVTAERKIGKKLNATI 64  
DB 29 ADEVKPGTYMDLTAKLQNPUNVRVL-----DLGGQNTTLPKEIEQL--KNLQKLY 77  
OY 65 LYKN-----KHLTDLKESHETTVMKAYVTTENGIR---YETQGT----- 102  
DB 78 LFDNRLLKTLPREIGQLNLQELNLSNQLTLPKEIGKLENLQRLDLDYDNRLLTLPKEIG 137  
OY 103 RTQTFEDVFVPSDYKNCVDYFVFKERGSDEGDEYELWVSDKIDKIP 148  
DB 138 KLQMLQTLVLSN---QTTLPRESGKLENLQELNLSNQLTLP 179

Search completed: August 1, 2003, 12:20:01  
Job time: 54.3467 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2003, 12:15:03 ; Search time 11.0933 Seconds

(without alignments)  
771.532 Million cell updates/sec

Title: US-10-087-195-8

Perfect score: 989

Sequence: 1 NPTWANEAKLGSYODAMKSL.....EKTVRNRYTDSCKPAPAQN 182

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	989	100.0	200	1 HBP2_RHIAP
2	331.5	33.5	190	1 HBP2_RHIAP
3	267.5	27.0	190	1 HBP1_RHIAP
4	99	10.0	200	1 RS4_THEVO
5	86	8.7	529	1 VSM6_TRYDB
6	86	8.7	692	1 2180_HUMAN
7	84	8.5	555	1 INV1_PEA
8	81	8.2	223	1 RB4A_ECHMU
9	81	8.2	610	1 C4BP_BOVIN
10	81	8.2	851	1 OPPE_MYCPN
11	81	8.2	2492	1 ATRX_HUMAN
12	80.5	8.1	1234	1 C9AH_MOUSE
13	80	8.1	246	1 PCN3_SULTO
14	80	8.1	1579	1 PEPI_YEAST
15	79.5	8.0	2184	1 POLG_ECO9B
16	79	8.0	405	1 NORF_NEIMB
17	79	8.0	580	1 SYO_XYLFM
18	79	8.0	1607	1 M3K4_HUMAN
19	78.5	7.9	212	1 Y331_MYCE
20	78.5	7.9	319	1 GCTA_ACIPE
21	78.5	7.9	2203	1 POLG_ECO9B
22	77.5	7.8	306	1 Y871_HAEIN
23	77.5	7.8	1305	1 RRPL_AHSY9
24	77	7.8	484	1 SYE_MYCCE
25	77	7.8	2034	1 FER1_GABEL
26	77	7.8	2245	1 MYSJ_DICDI
27	76.5	7.7	414	1 CDV1_MOUSE
28	76.5	7.7	451	1 UXAC_THEMA
29	76.5	7.7	522	1 NU62_HUMAN
30	76.5	7.7	522	1 NU62_RAT
31	76.5	7.7	526	1 NU62_MOUSE
32	76.5	7.7	1345	1 VCAP_HSV6U
33	76	7.7	397	1 YBP2_YEAST

34	76	7.7	2261	1 ABC1_MOUSE	P41233 mus musculus
35	75.5	7.6	662	1 MYR1_HUMAN	O33613 homo sapien
36	75.5	7.6	1536	1 Y984_THEMA	O9X079 thermotoga
37	75	7.6	405	1 NORF_NEIMB	O9JY43 neisseria m
38	75	7.6	415	1 ERK1_DICDI	P42523 dictyostel
39	75	7.6	448	1 RMUC_ZYMO	O9REQ3 zymomonas m
40	74.5	7.5	199	1 CN06_HUMAN	O8TAAT homo sapien
41	74.5	7.5	321	1 VG74_HSVSA	O01033 herpesvirus
42	74.5	7.5	1473	1 TOP2_ARATH	P30182 arabidopsis
43	74.5	7.5	1475	1 APU_THERY	P16950 t amylopall
44	74.5	7.5	1549	1 YIR3_YEAST	P40438 saccharomyc
45	74.5	7.5	1549	1 YJW2_YEAST	P40890 saccharomyc

## ALIGNMENTS

RESULT 1	HBPM_RHIAP	STANDARD	PRT	200 AA.
AC	077422:			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	Male-specific histamine-binding salivary protein precursor (MS-HBP) ..			
OS	Rhipicephalus appendiculatus (Brown ear tick).			
CC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;			
CC	Parasitiformes; Ixodida; Ixodidae; Ixodidae; Rhipicephalus.			
CC	NCBI Taxid=34631;			
CC	SEQUENCE FROM N.A.			
CC	TISSUE=Salivary gland;			
CC	MEDLINE=99288454; PubMed=10360182;			
CC	Paesen G.C., Adams P.L., Harlos K., Nuttall P.A., Stuart D.I.;			
CC	"Tick histamine-binding proteins: Isolation, cloning, and three-			
CC	dimensional structure."			
CC	Mol. Cell 3:661-671(1999).			
CC	-1- FUNCTION: BINDS HISTAMINE WITH A HIGH-AFFINITY. THE ABILITY TO			
CC	OUTCOMPETE HISTAMINE RECEPTORS INDICATES THAT ITS FUNCTION IS TO			
CC	SUPPRESS INFLAMMATION DURING BLOOD FEEDING.			
CC	-1- SUBCELLULAR LOCATION: Secreted.			
CC	-1- SUBCELLULAR LOCATION: Secreted.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@sib.ch">license@sib.ch</a> ).			
CC	EMBL: U96082; AAC63108.1; ..			
CC	HSSP: O77421; 10PT.			
CC	InterPro: IPR002970; His_binding.			
CC	Pfam: PF02098; His_binding; 1.			
CC	ProDom: PD152455; His_binding; 1.			
CC	Signal.			
CC	SIGNAL			
CC	CHAIN			
CC	1	18	POTENTIAL.	
CC	19	200	MALE-SPECIFIC HISTAMINE-BINDING SALIVARY	
CC	PROTEIN.			
CC	DISULFID			
CC	65	193	BY SIMILARITY.	
CC	DISULFID			
CC	137	169	BY SIMILARITY.	
CC	SEQUENCE			
CC	200 AA; 22851 MW; C46A1C8CB8CA008 CRC64;			
CC	Query Match			
CC	100.0%; Score 989; DB 1; Length 200;			
CC	Best Local Similarity 100.0%; Pred. No. 1.2e-78;			
CC	Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
CC	1 NPTWANEAKLGSYODAMKSLQODONKRYVLAQATQTDGVGGEFTCVSYAERKIGKKL 60			
CC				
CC	19 NPTWANEAKLGSYODAMKSLQODONKRYVLAQATQTDGVGGEFTCVSYAERKIGKKL 78			
CC				
CC	61 NATLLYKNKHLTDKESHETITVWKAAYDTTENGICYETQGTPTQTFEDVVFSDYKNC 120			
CC				

DB 79 NATIYKNNKHLTDKESHEITVWKAVDYTTENGIRKOTGTOTEDVFEVSDYKNC 138  
 QY 121 VIFVPERKSGDESDYELWVSEDKIDKIPDCKCFMAYFAOQOEKTVANNVYTDSSCKPAPA 180  
 DB 139 VIFVPERKSGDESDYELWVSEDKIDKIPDCKCFMAYFAOQOEKTVANNVYTDSSCKPAPA 198  
 QY 181 QN 182  
 DB 199 QN 200

RESULT 2  
 HBP2\_RH1AP STANDARD; PRT: 190 AA.  
 ID HBP2\_RH1AP  
 AC 077421;  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Female-specific histamine-binding protein 2 precursor (FS-HBP2).  
 OS Rhipicephalus appendiculatus (Brown ear tick).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Ixodidae; Rhipicephalus.  
 NCBI\_TaxID=34631;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (1.25 ANGSTROMS).  
 RC TISSUE=Salivary gland;  
 RX MEDLINE=9288454; PubMed=10360182;  
 RA Paesen G.C., Adams P.L., Harlos K., Nuttall P.A., Stuart D.I.;  
 RT "Tick histamine-binding proteins: Isolation, cloning, and three-dimensional structure."  
 RL Mol. Cell 3:661-671(1999).  
 CC -1- FUNCTION: BINDS HISTAMINE WITH A HIGH-AFFINITY. THE ABILITY TO OUTCOMPETE HISTAMINE RECEPTORS INDICATES THAT ITS FUNCTION IS TO SUPPRESS INFLAMMATION DURING BLOOD FEEDING.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: U96081; AAC63107.1; -;  
 DR PDB: 1QF7; 19-APR-00.  
 DR PDB: 1QFV; 19-APR-00.  
 DR InterPro: IPR002970; His\_binding.  
 DR Pfam: PF02098; His\_binding; 1.  
 DR Prodom: PD152455; His\_binding; 1.  
 KW Signal; 3D-structure.  
 FT SIGNAL 1 19  
 FT CHAIN 20 190  
 FT  
 FT DISULFID 67 188  
 FT DISULFID 138 167  
 FT TURN 23 24  
 FT HELIX 27 30  
 FT HELIX 31 33  
 FT HELIX 36 41  
 FT TURN 42 45  
 FT STRAND 48 53  
 FT STRAND 57 58  
 FT TURN 59 61  
 FT STRAND 62 63  
 FT TURN 64 64  
 FT STRAND 66 76  
 FT TURN 77 80  
 FT STRAND 81 88  
 FT TURN 91 92  
 FT STRAND 97 106  
 FT TURN 109 110  
 FT STRAND 116 121

FT TURN 122 123  
 FT STRAND 126 135  
 FT TURN 136 137  
 FT STRAND 138 143  
 FT STRAND 152 157  
 FT TURN 160 161  
 FT HELIX 165 174  
 FT TURN 175 177  
 FT STRAND 181 182  
 FT TURN 186 188  
 SQ SEQUENCE 190 AA; 21464 MW; 6923A3E902552B6F CRC64;

Query Match 33.58; Score 331.5; DB 1; Length 190;  
 Best Local Similarity 42.08; Pred. No. 7.5e-22;  
 Matches 74; Conservative 29; Mismatches 62; Indels 11; Gaps 7;

QY 2 PTVANEAKLGSDYDAMKSLQDQNKRRYLAAOTGTGCVGGEPTCVTAERIG--KKK 59  
 DB 22 PDWADENANGAHQDAMKSLADVENVYWKATYKNDPVWGNDFTCVGANADVNEDEKS 81  
 QY 60 LNATIIYKNNKHLTDKESHEITVWKAVDYTTENGIRKOTGTOTEDVFEVSDYKNC 119  
 DB 82 IQAEFFLEPMNADINMQATEKVTAVKMYGNRENAFRETED--GGVFTDVIAYSD-DNC 138  
 QY 120 DVIFVPERKSGDESDYELWVSEDKIDKIPDCKCFMAYFAOQOEKTVANNVYTDSSC 175  
 DB 139 DVIVFPGTDGNEEG-YELWTTD--YDNIIPANCLNKENEYAVGRE--TRDVFET-SAC 188

RESULT 3  
 HBP1\_RH1AP STANDARD; PRT: 190 AA.  
 ID HBP1\_RH1AP  
 AC 077420;  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Female-specific histamine-binding protein 1 precursor (FS-HBP1).  
 OS Rhipicephalus appendiculatus (Brown ear tick).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Ixodidae; Rhipicephalus.  
 NCBI\_TaxID=34631;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Salivary gland;  
 RX MEDLINE=9288454; PubMed=10360182;  
 RA Paesen G.C., Adams P.L., Harlos K., Nuttall P.A., Stuart D.I.;  
 RT "Tick histamine-binding proteins: Isolation, cloning, and three-dimensional structure."  
 RL Mol. Cell 3:661-671(1999).  
 CC -1- FUNCTION: BINDS HISTAMINE WITH A HIGH-AFFINITY. THE ABILITY TO OUTCOMPETE HISTAMINE RECEPTORS INDICATES THAT ITS FUNCTION IS TO SUPPRESS INFLAMMATION DURING BLOOD FEEDING.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: U96080; AAC63106.1; -;  
 DR HSSP: 077421; 1QF7.  
 DR InterPro: IPR002970; His\_binding.  
 DR Pfam: PF02098; His\_binding; 1.  
 DR Prodom: PD152455; His\_binding; 1.  
 KW Signal.  
 FT SIGNAL 1 18  
 FT CHAIN 19 190  
 FT  
 FT DISULFID 66 187  
 FT DISULFID 137 166







CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: X85327; CAA59677.1; -  
DR PIR: T06491; T06491.  
DR InterPro: IPR001362; Glyco\_hydro\_32.  
DR Pfam: PF00251; Glyco\_hydro\_32; 1.  
DR SMART: SM00640; Glyco\_32; 1.  
DR PROSITE: PS00609; GLYCOSYL\_HYDROL\_F32; 1.  
KW Hydrolase; Glycosidase; Glycoprotein; Signal.  
FT SIGNAL 1 22  
FT CHAIN 1 22  
FT SIGNAL 23 555  
FT CHAIN 1 22  
FT ACT\_SITE 140 140  
FT CARBOHD 154 154  
FT CARBOHD 181 181  
FT CAROHD 337 337  
SQ SEQUENCE 555 AA; 62655 MW; 7B548E251CFD5D44 CRC64;

Query Match  
Best Local Similarity 8.5%; Score 84; DB 1; Length 555;  
Matches 38; Conservative 22; Mismatches 52; Indels 50; Gaps 8;

OY 2 PTWANEAKLSGYODAKMSLQODONKRYLAQOTTTDGVGSEFTCVSTAEKICKKLN 61  
DB 174 PTANKINSSSEFRDPTTS-----WLGR-----DGFWR-----VLGSKIDTKGM- 212  
OY 62 ATILKKNHLDLKE-----SHETITWKAQDY-----TTENG---IKYE 98  
DB 213 -AIIKSKNFYDWVAKKPLHSAEGTGMECPDFVLDKLLRTGVDTSRNGDDVRRV 271  
OY 99 TCGTRTQTFEDVFEVSDYKNCVDIVFKERGSDEG-----DY 135  
DB 272 LKVSIDDPKHDHYLGSTYDVVQVDFPEKGFEDNFKVLRDY 313

RESULT 8  
RBA4\_ECHMU STANDARD: PRT; 223 AA.  
ID RBA4\_ECHMU STANDARD: PRT; 223 AA.  
AC 09G933: STANDARD: PRT; 223 AA.  
DT 15-SEP-2003 (Rel. 42, Created)  
DT 15-SEP-2003 (Rel. 42, Last sequence update)  
DE 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Probable Ras-related protein Rab-4A.  
OS Echinococcus multilocularis.  
OC Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;  
OC Cyclophyllidae; Taeniidae; Echinococcus.  
OX NCBI\_TaxID=6211;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H-95;  
RX MEDLINE=20556267; PubMed=10973970;  
RA Brehm K., Jensen K., Frosch M.;  
RA "mRNA trans-splicing in the human parasitic cestode Echinococcus  
RT multilocularis.";  
RL J. Biol. Chem. 275:38311-38318(2000).  
CC -I- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR  
CC TRAFFIC (BY SIMILARITY).  
CC -I- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.  
CC -I- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: AJ292375; CAC18549.1; -  
DR HSSP: P05713; 3RAB.  
DR InterPro: IPR003579; GTPase\_Rab.  
DR InterPro: IPR001806; Ras\_trnsmg.  
DR InterPro: IPR005225; Small\_GTP.  
DR Pfam: PF00071; ras; 1.  
DR PRINTS: PR00449; RASPRNSFRNMG.  
DR SMART: SM00175; RAB; 1.  
KW TIGRFRMS: TIGR00231; small\_GTP; 1.  
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.  
FT NP\_BIND 16 23  
FT NP\_BIND 64 68  
FT NP\_BIND 122 125  
FT DOMAIN 38 46  
FT DOMAIN 221 221  
FT LIPID 223 223  
SQ SEQUENCE 223 AA; 24454 MW; 9EB6D6166E2089A4 CRC64;

Query Match  
Best Local Similarity 8.2%; Score 81; DB 1; Length 223;  
Matches 22; Conservative 17; Mismatches 34; Indels 8; Gaps 3;

OY 64 ILKKNHLDLKEHETITWKAQDYTTENGIRY-ETGTRTQTFEDVFEVSDYKNCVY 122  
DB 118 ILVGNK--KDLRDTGQVTHWEANFPAQNGIQTETSLATGENIDAF-----SCVRV 170  
OY 123 FVFKERGSDEGEIYVWSEDK 143  
DB 171 LLSKSGELGDRLLVGSNK 191

RESULT 9  
C4BP\_BOVIN STANDARD: PRT; 610 AA.  
ID C4BP\_BOVIN STANDARD: PRT; 610 AA.  
AC Q28065:  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE C4b-binding protein alpha chain precursor (c4bp).  
GN C4BP.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=95015909; PubMed=7930621;  
RA Hillarp A., Thern A., Dahlbaeck B.;  
RA "Bovine C4b binding protein. Molecular cloning of the alpha- and  
RT beta-chains provides structural background for lack of complex  
RT formation with protein S.";  
RL J. Immunol. 153:4190-4199(1994).  
CC -I- FUNCTION: C4BP CONTROLS THE CLASSICAL PATHWAY OF COMPLEMENT  
CC ACTIVATION. IT BINDS AS A COFACTOR TO C3B/C4B INACTIVATOR  
CC (C3BIN), WHICH THEN HYDROLYZES THE COMPLEMENT FRAGMENT C4B. IT  
CC ALSO ACCELERATES THE DEGRADATION OF THE C4B2A COMPLEX (C3  
CC CONVERTASE) BY DISSOCIATING THE COMPLEMENT FRAGMENT C2A. ALPHA  
CC CHAIN BINDS C4B. IT INTERACTS ALSO WITH SERUM AMYLOID P COMPONENT.  
CC -I- SUBUNIT: DISULFIDE-LINKED COMPLEX OF C4BP ALPHA AND BETA CHAINS  
CC (BY SIMILARITY).  
CC -I- SIMILARITY: TO C4BP BETA CHAIN AND TO PIG APOLIPOPROTEIN R.  
CC -I- SIMILARITY: Contains 8 Sushi (SCR) domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC -----
DR EMBL: Z31693; CAAB3498.1;
DR PIR: I46001; I46001.
DR HSSP: P10998; IYVD.
DR InterPro: IPR000436; Sush1_SCR_CCP.
DR Pfam: PF00084; Sush1; 8.
DR SMART: SM00032; CCP; 8.
DR Complement pathway; Plasma; Glycoprotein; Repeat; Sush1; Signal.
KW SIGNAL
FT CHAIN 1 48
FT DOMAIN 49 610
FT DOMAIN 108 170
FT DOMAIN 111 170
FT DOMAIN 173 235
FT DOMAIN 238 295
FT DOMAIN 298 363
FT DOMAIN 366 426
FT DOMAIN 428 484
FT DOMAIN 486 542
FT DOMAIN 50 95
FT DISULFID 80 107
FT DISULFID 112 153
FT DISULFID 139 169
FT DISULFID 174 217
FT DISULFID 203 234
FT DISULFID 239 281
FT DISULFID 267 294
FT DISULFID 299 350
FT DISULFID 334 362
FT DISULFID 366 390
FT DISULFID 367 425
FT DISULFID 402 425
FT DISULFID 429 471
FT DISULFID 457 483
FT DISULFID 487 528
FT DISULFID 514 541
FT DISULFID 549 549
FT DISULFID 561 561
FT CARBOHYD 66 66
FT CARBOHYD 221 221
FT CARBOHYD 525 525
FT CARBOHYD 602 602
SQ SEQUENCE 610 AA; 68886 MW; D806B270E8A06B58 CRC64;

Query Match 8.28; Score 81; DB 1; Length 610;
Best Local Similarity 23.7%; Pred. No. 13;
Matches 32; Conservative 14; Mismatches 47; Indels 42; Gaps 5;

QY 78 HETITWKAYD-----YTTENGIKYETGCTRTGFEDVVFESDYKNCVIFPKRSGSD 131
DB 315 HQOTBQVYDIDGFLVLSYKHFQKPEPTDPTVTCOSNLNEMSPYIECKEYCCPEPLN 374
QY 132 EGYEL-----WSEDKID-----KIPDC---CKFTMA 156
DB 375 YGSITLHRRPSTHCHYISGDKISYCHSKYWFEDALCTKGTWSPRTPECRDCK-SP 433
QY 157 YFAQOQEKTVRNYT 171
DB 434 VIAHGKHVSKFT 448

RESULT 10
OPPF_MYCPN STANDARD; PRT; 851 AA.
AC P75551;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Oligopeptide transport ATP-binding protein oppf.
GN OPF OR MPN218 OR MP613.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_Taxid=2104;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 29342 / M129;
RX MEDLINE-97105885; PubMed-994633;
RA Himmelfreich R., Hilbert H., Plagens H., Pirkle E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae."
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- FUNCTION. PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC FOR OLIGOPEPTIDES. PROBABLY RESPONSIBLE FOR ENERGY COUPLING TO THE
CC TRANSPORT SYSTEM (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Membrane-associated (potential).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB000058; AAB96261.1;
DR PIR: S73939; S73939.
DR InterPro: IPR003593; AAA_ATPase.
DR Pfam: PF00005; ABC_tran; 1.
DR Prodom: PD000006; ABC_transporter; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE: PS50893; ABC_TRANSPORTER_2; 1.
KW Peptide transport; Transport; Membrane; ATP-binding;
FT NP_BIND 48 55
FT SEQUENCE 851 AA; 98611 MW; 6D3A3EC76B27DCC2 CRC64;

Query Match 8.28; Score 81; DB 1; Length 851;
Best Local Similarity 23.7%; Pred. No. 20;
Matches 44; Conservative 26; Mismatches 60; Indels 56; Gaps 8;

QY 3 TWANEAKLGSYDA-KKSLOO-DONKRYLAQNTOTTDGWSGEFVCSTAKIKKK- 59
DB 251 TIOQKAVSTQAFATWOKLQOTKONLKAFAQ-----MAEELONKPR 292
QY 60 --INATILYKXKHLTLKESHETITWKAYDYTTENGIKYETGCTRTGFEDVVFES--- 114
DB 293 IYLNAML-----TTKNVYKDSRG--NTQLDVFARFSYND 326
QY 115 --DYKNCVYFVKRSGDEGDELVSE--DKIDKIPDCKTMAVFAQOQEKTVRNY 170
DB 327 MVDKRRLLVLYSEYKALPEFYDNMTHONADRDELTVAVFEDLLIDVYIALNRPANVE 386
QY 171 TDSSCK 176
DB 387 SDAKAE 392

RESULT 11
ATRX_HUMAN
ID ATRX_HUMAN STANDARD; PRT; 2492 AA.
AC P46100; P51068; Q15886; Q9H021; Q9NTS3;
DT 01-NOV-1995 (Rel. 32, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Transcriptional regulator ATRX (X-linked helicase II) (X-linked
DE nuclear protein) (XNP) (znf-HX).
GN ATRX OR RAD54L OR XH2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4 AND 5), VARIANT SER-1860, AND

```

RP VARIANTS ATR-X.  
RX MEDLINE-97123494; PubMed-8968741;  
RA Picketts D.J., Higgs D.R., Bachoo S., Blake D.J., Quarrell O.W.J.,  
RA Gibbons R.J.;  
RT "ATR encodes a novel member of the SNF2 family of proteins: mutations  
RT point to a common mechanism underlying the ATR-X syndrome.";  
RL Hum. Mol. Genet. 5:1899-1907(1996).  
RN [12]  
RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 4).  
RX MEDLINE-97386582; PubMed-9244431.  
RA Villard L., Loefer A.-M., Cardoso C., Proud V., Chlaroni P.,  
RA Colleaux L., Schwartz C., Fontes M.;  
RT "Determination of the genomic structure of the XNP/ATR-X gene encoding  
RT a potential zinc finger helicase.";  
RL Genomics 43:149-155(1997).  
RN [13]  
RP SEQUENCE OF 860-2492 FROM N.A.  
RX MEDLINE-95179111; PubMed-7874112;  
RA Stayton C.L., Dabovic B., Gullisano M., Gecz J., Broccoli V.,  
RA Giovannazzi S., Bossolasco M., Monaco L., Raetan S., Boncinelli E.,  
RA Bianchi M.E., Gonzalez G.G.;  
RT "Cloning and characterization of a new human Xq13 gene, encoding a  
RT putative helicase.";  
RL Hum. Mol. Genet. 3:1957-1964(1994).  
RN [14]  
RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.  
RX MEDLINE-94214473; PubMed-8162050;  
RA Gecz J., Pollard H., Gonzalez G., Villard L., Stayton C.L.,  
RA Millaesau P., Khrestchatsky M., Fontes M.;  
RT "Cloning and expression of the murine homologue of a putative human  
RT X-linked nuclear protein gene closely linked to PK1 in Xq13.3.";  
RL Hum. Mol. Genet. 3:39-44(1994).  
RN [15]  
RP SEQUENCE OF 2401-2492 FROM N.A., AND VARIANTS ATR-X.  
RX MEDLINE-95211835; PubMed-7697714;  
RA Gibbons R.J., Picketts D.J., Villard L., Higgs D.R.;  
RT "Mutations in a putative global transcriptional regulator cause X-  
RT linked mental retardation with alpha-thalassemia (ATR-X syndrome).";  
RL Cell 80:837-845(1995).  
RN [16]  
RP SEQUENCE OF 1375-2492 FROM N.A.  
RA Pearce A., Chapman J.;  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [17]  
RP E2H2 BINDING.  
RX MEDLINE-98167853; PubMed-9499421;  
RA Cardoso C., Timait S., Villard L., Khrestchatsky M., Fontes M.,  
RA Colleaux L.;  
RT "Specific interaction between the XNP/ATR-X gene product and the SET  
RT domain of the human E2H2 protein.";  
RL Hum. Mol. Genet. 7:679-684(1998).  
RN [18]  
RP SUBCELLULAR LOCATION, AND ASSOCIATION WITH PERICENTROMERIC  
RP HETEROCHROMATIN.  
RX MEDLINE-20040663; PubMed-10570185;  
RA McDowell T.L., Gibbons R.J., Sutherland H., O'Rourke D.M.,  
RA Bickmore W.A., Pombo A., Turley H., Gatter K., Picketts D.J.,  
RA Buckle V.J., Chapman L., Rhodes D., Higgs D.R.;  
RT "Localization of a putative transcriptional regulator (ATRX) at  
RT pericentromeric heterochromatin and the short arms of acrocentric  
RT chromosomes.";  
RL Proc. Natl. Acad. Sci. U.S.A. 96:13983-13988(1999).  
RN [19]  
RP DISEASE.  
RX MEDLINE-20213147; PubMed-10751095;  
RA Villard L., Fontes M., Ades L.C., Gecz J.;  
RT "Identification of a mutation in the XNP/ATR-X gene in a family  
RT reported as Smith-Fineman-Myers syndrome.";  
RL Am. J. Med. Genet. 91:83-85(2000).  
RN [10]  
RP VARIANT ATR-X SER-1713.  
RX MEDLINE-97196774; PubMed-9043863;  
RA Villard L., Lacombe D., Fontes M.;

RT "A point mutation in the XNP gene, associated with an ATR-X phenotype  
RT without alpha-thalassemia.";  
RL Eur. J. Hum. Genet. 4:316-320(1996).  
RN [11]  
RP VARIANT JM GLN-2131.  
RX MEDLINE-96224392; PubMed-8630485;  
RA Villard L., Gecz J., Mattei J.-F., Fontes M., Saugier-Verber P.,  
RA Munnich A., Lyonnet S.;  
RT "XNP mutation in a large family with Juberger-Marsidi syndrome.";  
RL Nat. Genet. 12:359-360(1996).  
RN [12]  
RP VARIANTS ATR-X.  
RX MEDLINE-97467722; PubMed-9326931;  
RA Gibbons R.J., Bachoo S., Picketts D.J., Aftimos S., Azenbauer B.,  
RA Begoffen J., Berry S.A., Dahl N., Fryer A., Keppler K., Kurosawa K.,  
RA Levin M.L., Masuno M., Neil G., Pierpont M.E., Slaney S.F.,  
RA Higgs D.R.;  
RT "Mutations in transcriptional regulator ATRX establish the functional  
RT significance of a PHD-like domain.";  
RL Nat. Genet. 17:146-148(1997).  
RN [13]  
RP VARIANT ATR-X LEU-246.  
RX MEDLINE-20123062; PubMed-10660327;  
RA Fichera M., Romano C., Castiglia L., Falla P., Ruberto C., Amata S.,  
RA Greco D., Cardoso C., Fontes M., Ragusa A.;  
RT "New mutations in XNP/ATR-X gene: a further contribution to  
RT genotype/phenotype relationship in ATR/X syndrome.";  
RL Hum. Mutat. 12:214-214(1998).  
RN [14]  
RP VARIANT SHS LYS-1742.  
RX MEDLINE-99347960; PubMed-10417298;  
RA Loefer A.-M., Millan J.M., Villard L., Orellana C., Cardoso C.,  
RA Prieto F., Fontes M., Martinez F.;  
RT "Mutation of the XNP/ATR-X gene in a family with severe mental  
RT retardation, spastic paraplegia and skewed pattern of X inactivation  
RT demonstration that the mutation is involved in the inactivation  
RT bias.";  
RL Am. J. Hum. Genet. 65:558-562(1999).  
RN [15]  
RP VARIANT CMS THR-2050.  
RX MEDLINE-99326061; PubMed-10398237;  
RA Abidi F., Schwartz C.E., Carpenter N.J., Villard L., Fontes M.,  
RA Curtis M.;  
RT "Carpenter-Maziri syndrome results from a mutation in XNP.";  
RL Am. J. Med. Genet. 85:249-251(1999).  
RN [16]  
RP VARIANTS ATR-X GIU-175; 178-VAL-LYS-198 DEL; SER-190; PRO-219;  
RX LEU-246 AND CYS-249.  
RX MEDLINE-99219535; PubMed-10204841;  
RA Villard L., Bonlino M.-C., Abidi F., Ragusa A., Beloune J.,  
RA Loefer A.-M., Seaver A., Bonnefont J.-P., Romano C., Fichera M.,  
RA Lacombe D., Hanauer A., Philip N., Schwartz C.E., Fontes M.;  
RT "Evaluation of a mutation screening strategy for sporadic cases of  
RT ATR-X syndrome.";  
RL J. Med. Genet. 36:183-186(1999).  
RN [17]  
RP VARIANTS ATR-X SER-179; LEU-190; ILE-194; CYS-246; PHE-1552; SER-1645  
RX AND CYS-1847.  
RX MEDLINE-20451413; PubMed-10995512;  
RA Wada T., Kubota T., Fukushima Y., Saitoh S.;  
RT "Molecular genetic study of Japanese patients with X-linked alpha-  
RT thalassemia/mental retardation syndrome (ATR-X).";  
RL Am. J. Med. Genet. 94:242-248(2000).  
RN [18]  
RP FUNCTION: COULD BE A GLOBAL TRANSCRIPTIONAL REGULATOR. MODIFIES  
RP GENE EXPRESSION BY AFFECTING CHROMATIN. MAY BE INVOLVED IN BRAIN  
RP DEVELOPMENT AND FACIAL MORPHOGENESIS.  
CC -1- SUBUNIT: PROBABLY BINDS E2H2. BINDS ANNEXIN V IN A CALCIUM AND  
CC PHOSPHATIDYLCHOLINE/PHOSPHATIDYLSERINE-DEPENDENT MANNER (BY  
CC similarity).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH PERICENTROMERIC  
CC HETEROCHROMATIN DURING INTERPHASE AND MITOSIS, PROBABLY BY  
CC INTERACTING WITH HP1.  
CC -1- ALTERNATIVE PRODUCTS:

CC Event-Alternative splicing; Named isoforms=5;  
CC Name=4;  
CC IsoId=P46100-1; Sequence-Displayed;  
CC Name=1;  
CC IsoId=P46100-2; Sequence-VSP\_000575;  
CC Name=2;  
CC IsoId=P46100-3; Sequence-VSP\_000574;  
CC Name=3;  
CC IsoId=P46100-4; Sequence-VSP\_000576;  
CC Name=5;  
CC IsoId=P46100-5; Sequence-VSP\_000574, VSP\_000576;  
CC TISSUE SPECIFICITY: Ubiquitous;  
CC -1- DISEASE: Defects in ATRX are the cause of X-linked alpha-thalassemia/mental retardation syndrome (ATR-X) [MIM:301040]. ATR-X is an X-linked disorder comprising severe psychomotor retardation, facial dysmorphism, urogenital abnormalities, and alpha-thalassemia. An essential phenotypic trait are hemoglobin H erythrocyte inclusions.  
CC -1- DISEASE: Defects in ATRX are the cause of Sutherland-Haan X-linked mental retardation syndrome (SHS) [MIM:309470]. It is characterized by severe mental retardation with spastic paraplegia, microcephaly, short stature and cryptorchidism.  
CC -1- DISEASE: Defects in ATRX are a cause of Smith-Fineman-Wyers syndrome (SFM) [MIM:309580]. Clinical features include severe mental retardation, microcephaly, growth failure, facial anomalies and bilateral cryptorchidism. Due to the clinical overlap with ATR-X syndrome, some patients originally diagnosed as having SFM, might be affected by a variant of ATR-X syndrome which lack hemoglobin h inclusions.  
CC -1- DISEASE: Defects in ATRX are the cause of Carpenter-Wariri syndrome (CMS), an X-linked recessive condition characterized by moderate mental retardation, short stature, brachydactyly with excessive skin creases, and widening of the knuckles.  
CC -1- DISEASE: Defects in ATRX are the cause of Jubb-Marsidi syndrome (JM) [MIM:309590]. JM is a rare X-linked recessive disease characterized by severe mental retardation, growth failure, sensorineural deafness, microgenitalism and early death.  
CC -1- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY.  
CC -1- SIMILARITY: CONTAINS 1 PHD-TYPE ZINC FINGER.  
CC  
Query Match 8.2%; Score 81; DB 1; Length 2492;  
Best Local Similarity 26.4%; Pred. No. 71;  
Matches 42; Conservative 21; Mismatches 56; Indels 40; Gaps 8;  
OY 32 QATQTGTV--WGEFTCVSTAEKIGKKKLNATLLYKKNHLDKESHETTVMKAYD 88  
DB 913 QASASTDGVKLSGKEQSFSLVRYAETK-----EKSRL-----KTKCKKVD 959  
OY 89 YTTENGIKYEOTGTRQTFEDVVFSDYKNCQVIFPKREGSP-----EDDYELWVS 140  
DB 960 GLSDIAEKFLKKQSDSTESD-----DKQSKRGTEKKRPDSFKRKVIMEOOYE--SS 1012  
OY 141 EDKIDKIP---DCCKFTMAVFAOQOEKTVNRYNTDSSCK 176  
DB 1013 SDGTEKLPREIECHPEKGI-----KQIKNGTIDDEKK 1045  
RESULT 12  
CPAH MOUSE STANDARD; PRT; 1234 AA.  
ID CFAH MOUSE  
AC P06909;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Complement factor H precursor (Protein beta-1-H).  
GN HPL OR CFH  
OS Mus musculus (Mouse).  
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxId=10090;  
RN [1]  
RP SEQUENCE FROM N.A.

RX MEDLINE=66233353; PubMed=2940596;  
RA Kristensen T., Tack B.F.;  
RT "Murine protein H is comprised of 20 repeating units, 61 amino acids in length."  
RL Proc. Natl. Acad. Sci. U.S.A. 83:3963-3967(1986).  
RN [2]  
RP SEQUENCE OF 1-19 FROM N.A.  
RC STRAIN=BA1B/C;  
RX MEDLINE=90148935; PubMed=2533512;  
RA Munoz-Canoes P., Tack B.F., Vlk D.P.;  
RT "Analysis of complement factor H mRNA expression: dexmethasone and IFN-gamma increase the level of H in L cells."  
RL Biochemistry 28:9891-9897(1989).  
RN [3]  
RP SEQUENCE OF 1-18 FROM N.A.  
RX MEDLINE=90111033; PubMed=2136885;  
RA Natsume-Sakai S., Nonaka M., Nonaka M., Harada Y.N., Shreffler D.C., Moriaki K.;  
RT "Demonstration of an unusual allelic variation of mouse factor H by the complete cDNA sequence of the H.2 allotype."  
RL J. Immunol. 144:358-362(1990).  
CC -1- FUNCTION: Factor H functions as a cofactor in the inactivation of C3b by factor I and also increases the rate of dissociation of the C3b complex (C3 convertase) and the (C3b)NBB complex (C5 convertase) in the alternative complement pathway (by similarity).  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Synthesized by liver and secreted in plasma.  
CC -1- POLYMORPHISM: TWO CODOMINANT ALLELES OF FACTOR H ARE PRESENT IN MICE.  
CC  
CC -1- SIMILARITY: Contains 20 Sushi (SCR) domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; M12660; AAA37759.1; -;  
DR EMBL; J02891; AAA37795.1; -;  
DR EMBL; M31979; AAA37762.1; -;  
DR PIR; A26154; NEMSH.  
DR HSSP; P08603; IHFI.  
DR MGD; MGI:88385; CfH.  
DR InterPro; IPR000436; Sushi\_SCR\_CCP.  
DR Pfam; PF00084; sushi; 20.  
DR SMART; SM00032; CCP; 20.  
KW Complement alternate pathway; Plasma; Glycoprotein; Repeat; Sushi; Signal.  
FT 1 18 BY SIMILARITY.  
FT CHAIN 19 1234 COMPLEMENT FACTOR H.  
FT DOMAIN 20 81 SUSHI 1.  
FT DOMAIN 84 142 SUSHI 2.  
FT DOMAIN 145 206 SUSHI 3.  
FT DOMAIN 209 263 SUSHI 4.  
FT DOMAIN 266 321 SUSHI 5.  
FT DOMAIN 324 386 SUSHI 6.  
FT DOMAIN 388 443 SUSHI 7.  
FT DOMAIN 447 506 SUSHI 8.  
FT DOMAIN 508 565 SUSHI 9.  
FT DOMAIN 568 623 SUSHI 10.  
FT DOMAIN 628 684 SUSHI 11.  
FT DOMAIN 689 744 SUSHI 12.  
FT DOMAIN 751 803 SUSHI 13.  
FT DOMAIN 807 862 SUSHI 14.  
FT DOMAIN 866 932 SUSHI 15.  
FT DOMAIN 935 990 SUSHI 16.  
FT DOMAIN 993 1049 SUSHI 17.  
FT DOMAIN 1052 1108 SUSHI 18.  
FT DOMAIN 1113 1169 SUSHI 19.  
FT DOMAIN 1171 1234 SUSHI 20.  
FT DISULFID 21 66 BY SIMILARITY.

```

FT DISULFID 52 80 BY SIMILARITY.
FT DISULFID 85 129 BY SIMILARITY.
FT DISULFID 114 141 BY SIMILARITY.
FT DISULFID 146 192 BY SIMILARITY.
FT DISULFID 178 205 BY SIMILARITY.
FT DISULFID 210 251 BY SIMILARITY.
FT DISULFID 237 262 BY SIMILARITY.
FT DISULFID 267 309 BY SIMILARITY.
FT DISULFID 294 320 BY SIMILARITY.
FT DISULFID 325 374 BY SIMILARITY.
FT DISULFID 357 385 BY SIMILARITY.
FT DISULFID 389 431 BY SIMILARITY.
FT DISULFID 416 442 BY SIMILARITY.
FT DISULFID 448 494 BY SIMILARITY.
FT DISULFID 477 505 BY SIMILARITY.
FT DISULFID 509 553 BY SIMILARITY.
FT DISULFID 536 564 BY SIMILARITY.
FT DISULFID 569 610 BY SIMILARITY.
FT DISULFID 597 622 BY SIMILARITY.
FT DISULFID 629 672 BY SIMILARITY.
FT DISULFID 658 683 BY SIMILARITY.
FT DISULFID 690 732 BY SIMILARITY.
FT DISULFID 718 743 BY SIMILARITY.
FT DISULFID 752 791 BY SIMILARITY.
FT DISULFID 780 802 BY SIMILARITY.
FT DISULFID 808 850 BY SIMILARITY.
FT DISULFID 836 861 BY SIMILARITY.
FT DISULFID 867 920 BY SIMILARITY.
FT DISULFID 906 931 BY SIMILARITY.
FT DISULFID 936 978 BY SIMILARITY.
FT DISULFID 964 989 BY SIMILARITY.
FT DISULFID 994 1037 BY SIMILARITY.
FT DISULFID 1023 1048 BY SIMILARITY.
FT DISULFID 1053 1096 BY SIMILARITY.
FT DISULFID 1082 1107 BY SIMILARITY.
FT DISULFID 1114 1157 BY SIMILARITY.
FT DISULFID 1143 1168 BY SIMILARITY.
FT DISULFID 1172 1223 BY SIMILARITY.
FT DISULFID 1206 1233 BY SIMILARITY.
FT CARBOHYD 676 676 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 721 721 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 773 773 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 801 801 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1030 1030 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1061 1061 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1225 1225 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1234 AA; 139082 MM; C5AC02F341B95F77 CRC64;

```

Query Match 8.1%; Score 80.5; DB 1; Length 1234;  
 Best Local Similarity 29.8%; Pred. No. 34;  
 Matches 25; Conservative 13; Mismatches 19; Indels 27; Gaps 5;

```

OY 24 QNKRYLAQATQT---IDGVGGEFTCVSTAERKIGKKKLATLLYKKHLTDLKESHER 80
DB 1144 QOCQRYLLKGRKRTTTCNGKMSSEPTC-----LHACYIPEN-----IMESHNI 1184
OY 81 ITVM-----KAYDYTTEN---GIRY 97
DB 1185 ILKWRHTEKITYSHSGEDIEFGCKY 1208

```

RESULT 13  
 PCN3\_SULTO STANDARD: PRT: 246 AA.  
 AC 0973F5.  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE DNA polymerase sliding clamp C (Proliferating cell nuclear antigen homolog C) (PCNA C).  
 DE PCNC OR ST0944.  
 GN Sulfolobus tokodaii.  
 OS Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;

```

CC Sulfolobus.
OX NCBI_TaxID=119955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RC MEDLINE=21456156; PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankel A., Kousugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Koshida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermocacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001).
CC -I- FUNCTION: Sliding clamp subunit. Responsible for tethering the
CC catalytic subunit of DNA polymerase to DNA during high-speed
CC replication (By similarity).
CC -I- SUBUNIT: Homotrimer (By similarity).
CC -I- SIMILARITY: BELONGS TO THE PCNA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF000984; BAB65958.1; -
DR HAMAP; MF_00317; -; 1.
DR InterPro; IPR000730; Pr_cel_nuc_antlg.
DR Pfam; PF00705; PCNA.1
DR PRINTS; PR00339; PCNACTCLIN.
DR ProDom; PD002673; Pr_cel_nuc_antlg; 1.
DR PROSITE; PS01251; PCNA.1; FALSE_NEG.
DR PROSITE; PS00293; PCNA.2; FALSE_NEG.
DR DNA-binding; DNA replication; Complete proteome.
SQ SEQUENCE 246 AA; 27436 MM; BA370A5D22784E7F CRC64;

```

Query Match 8.1%; Score 80; DB 1; Length 246;  
 Best Local Similarity 26.2%; Pred. No. 5.6;  
 Matches 33; Conservative 17; Mismatches 38; Indels 38; Gaps 5;

```

OY 9 KLSGYQAMKS-----LQDDONKRYLLAQATQTSVGGEFTCVSTAERKIGKKKL 60
DB 70 KLEDFTDLVKTVTKNDSLVLFTDENQIKV-----TLDGVYERTFTPPSIASELETPNL 124
OY 61 NATLLYKKHL-----TDLKESHERI---TWKKA-----YDYTTENGI 95
DB 125 NLEFPFKKALTVFTFDIIDLIEDIGDSITFKMGCKLYLSANSDMGSSITELSTENG 184
OY 96 KYETFG 101
DB 185 LLESEG 190

```

RESULT 14  
 PEPL\_YEAST STANDARD: PRT: 1579 AA.  
 AC P32319.  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Vacuolar protein sorting/targeting protein PEPI precursor (Vacuolar carboxypeptidase sorting receptor VPS10) (Carboxypeptidase Y receptor) (CPY receptor).  
 DE PEPI OR VPS10 OR YBL017C OR YBL0302 OR YBL03.22.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C;  
 RX MEDLINE-93070614; PubMed-1441754;  
 RA van Dyck L., Purnelle B., Skala J., Goffeau A.;  
 RT "An 11.4 kb DNA segment on the left arm of yeast chromosome II  
 RT carries the carboxypeptidase Y sorting gene PEPI, as well as ACH1,  
 RT FUS3 and a putative ARS.";  
 RL Yeast 8:769-776(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-94243924; PubMed-8187177;  
 RA Marcusson E.G., Horadzovsky B.F., Cereghino J.L., Gharkhanian E.,  
 RA Ent S.D.;  
 RT "The sorting receptor for yeast vacuolar carboxypeptidase Y is  
 RT encoded by the VPS10 gene.";  
 RL Cell 77:579-586(1994).  
 CC -1- FUNCTION: INVOLVED IN THE INTRACELLULAR SORTING AND DELIVERY OF  
 CC SOLUBLE VACUOLAR PROTEINS. SEEMS TO SPECIFICALLY INTERACT WITH  
 CC CARBOXYPEPTIDASE Y (CPY). MAY EXECUTE MULTIPLE ROUNDS OF SORTING  
 CC BY CYCLING BETWEEN THE LATE GOLGI AND A PREVACUOLAR ENDOSOME-LIKE  
 CC COMPARTMENT. BINDS THE GOLGI-MODIFIED P2 FORM OF CPY, AND THIS  
 CC INTERACTION IS DEPENDENT ON THE PRESENCE OF AN INTACT CPY VACUOLAR  
 CC PROTEIN SORTING SIGNAL.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. LATE GOLGI  
 CC COMPARTMENT.  
 CC -1- DOMAIN: THE LUMENAL DOMAIN CONTAINS TWO REGIONS OF APPROXIMATELY  
 CC 650 AA THAT EXHIBIT 20% IDENTITY. THE CYTOPLASMIC DOMAIN MAY SERVE  
 CC AS A GOLDI RETENTION/RECYCLING SIGNAL.  
 CC -1- SIMILARITY: BELONGS TO THE PEPI FAMILY OF MEMBRANE GLYCOPROTEINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>  
 CC or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).  
 CC -----  
 DR EMBL: X68577; CAA48568.1; -;  
 DR EMBL: Z35778; CAA84836.1; -;  
 DR EMBL: U07621; AAA18831.1; -;  
 DR PIR: S25329; S25329.  
 DR SGD: S0000113; PEPI.  
 DR GO: GO:0006653; P:protein-vacuolar targeting; IMP.  
 DR GO: GO:0007034; P:vacuolar transport; IMP.  
 DR InterPro: IPR002860; GH\_BNR.  
 DR Pfam: PF02012; BNR; 13.  
 DR SMART: SM00602; VPS10; 2.  
 KW Transmembrane; Glycoprotein; Receptor; Signal; ATP-binding; Repeat.  
 FT SIGNAL 1 21  
 FT CHAIN 22 1579  
 FT  
 FT DOMAIN 22 1397  
 FT TRANSMEM 1398 1414  
 FT DOMAIN 1415 1579  
 FT NP\_BIND 180 187  
 FT CARBOHYD 96 96  
 FT CARBOHYD 170 170  
 FT CARBOHYD 447 447  
 FT CARBOHYD 793 793  
 FT CARBOHYD 1010 1010  
 FT CARBOHYD 1303 1303  
 FT CARBOHYD 19 19  
 FT CARBOHYD 38 38  
 FT CARBOHYD 54 54  
 FT CARBOHYD 74 74  
 FT CARBOHYD 94 94  
 FT CARBOHYD 119 126  
 FT CARBOHYD 152 153  
 FT CARBOHYD 165 167  
 FT CARBOHYD 202 202  
 FT CARBOHYD 214 214  
 S -> G (IN REF. 2).  
 S -> G (IN REF. 2).

FT CONFLICT 283 283 L -> I (IN REF. 2).  
 FT CONFLICT 378 378 F -> L (IN REF. 2).  
 FT CONFLICT 410 410 S -> N (IN REF. 2).  
 FT CONFLICT 426 426 K -> V (IN REF. 2).  
 FT CONFLICT 774 774 M -> R (IN REF. 2).  
 FT CONFLICT 1089 1089 S -> F (IN REF. 2).  
 FT CONFLICT 1266 1266 G -> S (IN REF. 2).  
 FT CONFLICT 1476 1476 A -> V (IN REF. 2).  
 FT CONFLICT 1557 1557 T -> A (IN REF. 2).  
 SQ SEQUENCE 1579 AA; 177776 MW; C778A0400612ECFA CRC64;  
 Query Match 8.1%; Score 80; DB 1; Length 1579;  
 Best Local Similarity 27.8%; Pred. No. 50;  
 Matches 35; Conservative 12; Mismatches 37; Indels 42; Gaps 7;  
 QY 84 WKADYTTENGIRK-----ENQGRTRQTFEDVFV-----SDYKNCQVIFPKR 127  
 Db 607 WNLAEGRVGVKTKIRKRODACLKVKFEDIQLEFETCDKCTEADYE-CAFEYRDA 665  
 QY 128 RGSDEGDYELWSEDKIDK-----IP-----DCKFTMAVFAQQQKRTVRYVYTD 173  
 Db 666 TGKCVDPYNYLIVSDVCDKTKKTVPVKPLQLVKGDKCKPM-----TVKSY--DI 714  
 QY 174 SCKPAP 179  
 Db 715 SCRGVP 720  
 RESULT 15  
 ID POLG\_EC01F STANDARD; PRT; 2184 AA.  
 AC 091734; Q66795; Q9Y1D6;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Genome polyprotein [containing: Coat protein VP4 (P1A); Coat protein VP2  
 DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Picornain 2A  
 DE (EC 3.4.22.29) (P2A); Core protein P2B; Core protein P2C; Core protein  
 DE P2A; genome-linked protein VP6 (P3B); Picornain 3C (EC 3.4.22.28)  
 DE (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D).  
 OS Echovirus 1 (strain Farouk / ATCC VR-1038).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Enterovirus.  
 OX NCBI\_TaxID=103908;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bergelson J.M.;  
 RT "Receptor interactions, infectious CDNA, and nucleotide sequences of  
 RT echovirus 1/8." to the EMBL/GenBank/DBJ databases.  
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 69-330 FROM N.A.  
 RX MEDLINE-96203959; PubMed-8627260;  
 RA Huttunen P., Santti J., Pulli T., Hyypiae T.;  
 RT "The major echovirus group is genetically coherent and related to  
 RT coxsackie B viruses.";  
 RL J. Gen. Virol. 77:715-725(1996).  
 RN [3]  
 RP SEQUENCE OF 570-853 FROM N.A.  
 RX MEDLINE-99138973; PubMed-9971773;  
 RA Oberste M.S., Maher K., Kilpatrick D.R., Pallansch M.A.;  
 RT "Molecular evolution of the human enteroviruses: correlation of  
 RT serotype with VP1 sequence and application to picornavirus  
 RT classification.";  
 RL J. Virol. 73:1941-1948(1999).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (3.55 ANGSTROMS) OF 2-850.  
 RX MEDLINE-99192672; PubMed-10089503;  
 RA Filman D.J., Wien M.W., Cunningham J.A., Bergelson J.M., Hogle J.M.;  
 RT "Structure determination of echovirus 1.";  
 RL Acta Crystallogr. D 54:1261-1272(1998).  
 CC -1- FUNCTION: P2A AND THE P3C POLYPEPTIDES ARE PROPEASES THAT CLEAVE  
 CC AT CERTAIN Q/G SITES IN THE POLYPEPTIDE. THEY ARE CYSTEINE



CC PROTEASES.  
CC -1- FUNCTION: It is thought that the P2C protein attaches to vesicular  
CC membranes and is associated with viral RNA synthesis.  
CC -1- CATALYTIC ACTIVITY: Selective cleavage of Gln-I-gly bond in the  
CC poliovirus polyprotein. In other picornavirus reactions Glu may be  
CC substituted for Gln, and Ser or Thr for Gly.  
CC -1- CATALYTIC ACTIVITY: Selective cleavage of Tyr-I-gly bond in the  
CC picornavirus polyprotein. In other picornavirus reactions Glu may  
CC be substituted for Gln, and Ser or Thr for Gly.  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +  
CC [RNA](N).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC CLEAVAGE BETWEEN VP4 AND VP2 IS AUTOCATALYTIC. VP1/VP2 IS  
CC CATALYZED BY P2A; ALL OTHER CLEAVAGES ARE CATALYZED BY P3C.  
CC -1- SIMILARITY: P2A PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC -1- SIMILARITY: P3C PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AF029859; AAC63944.2; -  
CC EMBL: X89531; CA61710.1; -  
CC EMBL: AF081314; AAD17718.1; -  
CC PDB: 1EV1, 27-JAN-99.  
CC InterPro: IPR003593; AAA\_ATPase.  
CC InterPro: IPR004004; Callic\_pol\_hel.  
CC InterPro: IPR000199; Cys\_protease\_3C.  
CC InterPro: IPR003138; Pico\_P1A.  
CC InterPro: IPR000081; Pico\_P2A.  
CC InterPro: IPR002527; Pico\_P2B.  
CC InterPro: IPR001676; RNv.  
CC InterPro: IPR000605; RNA\_helicase.  
CC InterPro: IPR007095; RNA\_pol\_DS\_PS.  
CC InterPro: IPR001205; RNA\_pol\_P3D.  
CC InterPro: IPR007094; RNA\_pol\_PSVir.  
CC Pfam: PF00548; Cys-Protease-3C; 1.  
CC Pfam: PF02226; Pico\_P1A; 1.  
CC Pfam: PF00947; Pico\_P2A; 1.  
CC Pfam: PF01552; Pico\_P2B; 1.  
CC Pfam: PF00073; RNv; 3.  
CC Pfam: PF00680; RNA\_dep\_RNA\_pol; 1.  
CC Pfam: PF00910; RNA\_helicase; 1.  
CC PRINTS: PR00918; CALICVIRUSNS.  
CC ProDom: PD001125; Cys\_protease\_3C; 1.  
CC ProDom: PD001306; Pico\_P2A; 1.  
CC ProDom: PD001274; Pico\_P2B; 1.  
CC SMART: SM00382; AAA; 1.  
CC Polyprotein; Coat protein; Core protein; Transferase; Myristate;  
CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.  
CC -----  
CC CHAIN 2 69 COAT PROTEIN VP4.  
CC CHAIN 70 330 COAT PROTEIN VP2.  
CC CHAIN 331 569 COAT PROTEIN VP3.  
CC CHAIN 570 850 COAT PROTEIN VP1.  
CC CHAIN 851 1000 PICORNAIN 2A.  
CC CHAIN 1001 1099 CORE PROTEIN P2B.  
CC CHAIN 1100 1428 CORE PROTEIN P2C.  
CC CHAIN 1429 1517 CORE PROTEIN P3A.  
CC CHAIN 1518 1539 GENOME-LINKED PROTEIN VPG.  
CC CHAIN 1540 1722 PICORNAIN 3C.  
CC CHAIN 1723 2184 RNA-DIRECTED RNA POLYMERASE.  
CC LIPID 2 MYRISTATE.  
CC ACT\_SITE 1686 1686 PROTEASE 3C (POTENTIAL).  
CC ACT\_SITE 1700 1700 PROTEASE 3C (POTENTIAL).  
CC CONFLICT 69 69 N -> M (IN REF. 2).  
CC CONFLICT 238 238 A -> R (IN REF. 2).

FT CONFLICT 611 611 A -> V (IN REF. 3).  
FT CONFLICT 645 645 Y -> F (IN REF. 3).  
FT CONFLICT 724 724 V -> I (IN REF. 3).  
FT CONFLICT 774 774 A -> T (IN REF. 3).  
FT STRAND 4 7  
FT STRAND 26 29  
FT HELIX 36 38  
FT STRAND 45 45  
FT TURN 50 50  
FT HELIX 51 54  
FT STRAND 57 57  
FT TURN 63 64  
FT TURN 80 81  
FT STRAND 91 94  
FT STRAND 93 86  
FT STRAND 101 102  
FT STRAND 103 105  
FT HELIX 103 105  
FT TURN 113 115  
FT STRAND 123 123  
FT HELIX 126 128  
FT TURN 129 129  
FT STRAND 133 134  
FT STRAND 138 140  
FT TURN 142 143  
FT STRAND 147 151  
FT TURN 152 152  
FT HELIX 153 157  
FT STRAND 159 160  
FT STRAND 168 180  
FT TURN 185 186  
FT STRAND 188 197  
FT TURN 198 198  
FT STRAND 203 203  
FT STRAND 206 207  
FT HELIX 212 214  
FT TURN 215 215  
FT STRAND 217 217  
FT TURN 218 219  
FT STRAND 220 220  
FT STRAND 222 223  
FT STRAND 223 223  
FT TURN 232 233  
FT STRAND 235 235  
FT STRAND 238 238  
FT HELIX 239 241  
FT TURN 242 245  
FT STRAND 246 246  
FT HELIX 248 253  
FT STRAND 256 260  
FT TURN 261 263  
FT STRAND 266 271  
FT STRAND 280 280  
FT TURN 282 284  
FT STRAND 285 285  
FT TURN 302 303  
FT STRAND 308 324  
FT TURN 338 341  
FT TURN 345 346  
FT STRAND 353 353  
FT TURN 356 357  
FT STRAND 369 370  
FT STRAND 372 372  
FT TURN 373 373  
FT HELIX 374 377  
FT TURN 378 378  
FT STRAND 381 382  
FT TURN 390 393  
FT TURN 394 394  
FT HELIX 397 399  
FT STRAND 400 405  
FT TURN 406 406  
FT TURN 410 410

FT	STRAND	412	417
FT	TURN	420	422
FT	HELIX	424	427
FT	TURN	428	428
FT	HELIX	430	435
FT	TURN	436	437
FT	STRAND	438	442
FT	STRAND	445	452
FT	TURN	455	456
FT	STRAND	458	458
FT	STRAND	460	466

Query Match	8.08;	Score 79.5;	DB 1;	Length 2184;
Best Local Similarity	32.18;	Prod No 87;		

```

QY      5 ANEKILSYDQAMKSLOODOKRRYYLQAOTOTTOGVNGEERT-----CVSTJARKICKKK 59
Db      1068 ATTLALIGCTSSPFWMLKOKVSOYGYGIPRAEKQNNQ--WIKKFTENTNCKGH--EWIALIKI 1124

QY      60 LNAATILYKNNKHLTLKESHERITPWAKAVDYTTENGIKYE----TQGTQTQTFEDVFVFSID 115
Db      1125 QKFTIEMLKVKRLPEVKEKHEFLNMLKQLPLEESQIATIEGSAPOSGODEQLFSNVQYFAH 1184

QY      116 YKNCDAIVFVERKESDSDEGYELWVSEDKIDKIPDCKEFTMAVFAQQQ----EKTVRN-V 169
Db      1185 Y-----CRKYAPLPYAAEAKRVFSLFKKMSNYI 1211

QY      170 YTDSSCKPAP 179
Db      1212 QFKSKRIEP 1221

```

```
Search completed: August 1, 2003, 12:15:53
Job time : 13.0933 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: August 1, 2003, 12:15:04 ; Search time 13.52 Seconds

(without alignments)  
1294.578 Million cell updates/sec

Title: US-10-087-195-8

Perfect score: 989

Sequence: 1 NPTWANEAKIGSYQDAWKS.....EKTVRNYTDSCKPAPQN 182

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_76:\*

1: p1r1:\*

2: p1r2:\*

3: p1r3:\*

4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	113.5	11.5	330	2 S70127	hypothetical prote
2	87	8.8	483	2 AB1247	two-component sens
3	86	8.8	483	2 AFI609	variant surface gl
4	86	8.7	529	2 S18453	myoblast city prot
5	85.5	8.6	1970	2 T03284	hypothetical prote
6	85	8.6	1959	2 AG1085	beta-fructofuranos
7	84	8.5	555	2 T06491	beta-fructofuranos
8	84	8.5	582	2 T06380	RNA-directed RNA p
9	83.5	8.4	1078	2 T00494	hypothetical prote
10	82.5	8.3	306	2 T12821	hypothetical prote
11	82.5	8.3	1802	2 G71616	C4b-binding protei
12	81	8.2	610	1 I46001	pol polypeptide
13	81	8.2	656	2 S30484	pol polypeptide
14	81	8.2	851	2 S73939	trans
15	81	8.2	1641	2 T18614	helicase II - huma
16	80.5	8.1	721	2 AFI254	penicillin-binding
17	80.5	8.1	1234	1 NBMSH	complement factor
18	80	8.1	641	2 PH1919	FL-160-4 protein -
19	80	8.1	1579	2 S25329	carboxypeptidase y
20	79	8.0	338	2 B75154	hypothetical prote
21	79	8.0	345	2 C89882	conserved hypotet
22	79	8.0	375	2 AC1350	N-acetylmuramoyl-L
23	79	8.0	405	2 G81184	late embryogenesis
24	79	8.0	479	2 T47561	glutaminyl-tRNA sy
25	79	8.0	580	2 G82693	hypothetical prote
26	79	8.0	1405	2 T04426	MAP kinase kinase
27	79	8.0	1607	2 T03022	hypothetical prote
28	78.5	7.9	212	2 F64236	glutamate COA-tr
29	78.5	7.9	320	2 S51051	

30	77.5	7.8	306	2 C64015	hypothetical prote
31	77.5	7.8	423	2 AB1142	N-carbamyl-L-amino
32	77.5	7.8	769	2 P89870	serine proteinase
33	77.5	7.8	1156	2 B70356	chromosome assembl
34	77	7.8	231	2 T41566	translin associate
35	77	7.8	334	2 T47983	NM-like protein -
36	77	7.8	421	2 C70038	maltose/maltodextr
37	77	7.8	484	2 A64251	glutamate-tRNA 119
38	77	7.8	656	2 S30483	pol polypeptide
39	77	7.8	719	2 S44237	trypanastigote sur
40	77	7.8	762	2 JC7174	N,N-dimethylformam
41	77	7.8	1108	2 T38673	probable transcrip
42	77	7.8	2034	2 T22147	hypothetical prote
43	77	7.8	2245	2 T18278	myosin heavy chain
44	76.5	7.7	414	2 AFI500	N-carbamyl-L-amino
45	76.5	7.7	434	2 C82885	membrane nuclease

## ALIGNMENTS

## RESULT 1

S70127

hypothetical protein YDR267C - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein D9954.11

C:Species: Saccharomyces cerevisiae

C:Date: 24-Aug-1996 #sequence\_revision 06-Sep-1996 #text\_change 19-Apr-2002

C:Accession: S70127

R:Le, T.

Submitted to the EMBL Data Library, May 1996

A:Description: The sequence of S. cerevisiae cosmid 9954.

A:Reference number: S70124

A:Accession: S70127

A:Molecule type: DNA

A:Residues: 1-330 &lt;LET&gt;

A:Cross-references: EMBL:U51030; NID:G1332633; PID:G1230640; GSPDB:GN00004; MIPS:YDR2

C:Genetics:

A:Gene: MIPS:YDR267C

A:Cross-references: SGD:S0002675

A:Map position: 4R

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

F:54-87/Domain: WD repeat homology &lt;WD1&gt;

F:103-136/Domain: WD repeat homology &lt;WD2&gt;

F:149-182/Domain: WD repeat homology &lt;WD3&gt;

Query Match

Best Local Similarity 11.5%; Score 113.5; DB 2; Length 330;

Matches 38; Conservative 17; Mismatches 62; Indels 21; Gaps 5;

Db

16 AMKSLQDQNRKRYLAQTOTDGVW-----GEFFCVSVAEKIKKLNATILYKN 68

114 AM-----SNDGYLATCSR-DKSWIMETDESGEYICISLVDEH--SDQVKHYIMPS 164

69 KHLTDLKESHTITVWKAVDYTE---NRIKYETGTRQTEDEVFSDYKNCDYIF 123

165 EALLASSYDDTVIRIMKDDDMCEAVLNGHEGVWSSDRDKEGVRLCSGSDSYVR 224

124 VPKERGSDEGDYELWVSE 141

225 VMKYMGGDEDDQDEWVCE 242

Db

225 VMKYMGGDEDDQDEWVCE 242

RESULT 2

AB1247

two-component sensor histidine kinase 11sk [imported] - Listeria monocytogenes (strai

C:Species: Listeria monocytogenes

C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001

C:Accession: AB1247

R:Glaser, P.; Frangou, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkut, G.; Madueno, E.; Maltournam, A.;

Ok, C.; Schuete, T.; Simoes, N.; Tierre, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland  
A:Title: Comparative genomics of *Listeria* species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AB1247  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-483 <GLA>  
A:Cross-references: GB:NC\_003210; PIDN:CAC9456.1; PID:g16410794; GSEDB:GN00177  
A:Experimental source: strain Ecd-e  
C:Genetics:  
A:Gene: lisk

```

Query Match      8.8%  Score 87;  DB 2;  Length 483;
Best Local Similarity 24.0%  Pred. No. 5.7;
Matches 37;  Conservative 29;  Mismatches 58;  Indels 30;  Gaps 7;

OY 18 KSLQODONKRRYYLAQATQTTDGVWGEETCVSYAEKIGKKRLNATILYKRNKHTDLKES 77
   |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 308 KKLVOEMDLISRADISQTK-----LQTTDVNAIVEGY---RRNPEVWENFITT-LKED 359
   |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

OY 78 HETTVKAKDY-----TTENGKIKETQCTRTQTEEDVFSVDYKNCVDYIYVPKRGRS 130
   |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 360 DTIDRLALQNHMLEQILLIIMDNAVKVSGDTEV---DMHYEKQOKOIH----- 406
   |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

OY 131 DEGDYELWVSEDKIDKIPDCKCFMMAVFAOOQEK 164
   ||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 407 DVRDYGGISOEELDKIFN--RFRVADKAKRSEK 438
   ||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

```

RESULT 3  
AF1609  
two-component sensor histidine kinase lisk [imported] - *Listeria innocua* (strain Cl1p11262)  
C:Species: *Listeria innocua*  
C:Date: 27-Nov-2001 #sequence.revision 27-Nov-2001 #text.change 27-Nov-2001  
C:Accession: AF1609  
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl, H. D.; Jones, L.M.; Karst, U. Science 294, 849-852, 2001  
A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maitournam, A.; Mekok, C.; Schlueder, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A. Title: Comparative genomics of *Listeria* species.  
A:Reference number: AB1077; MUID: 21537279; PMID: 11679669  
A:Accession: AF1609  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-483 <GIA>  
A:Cross-references: GB:AL592022; PIDN:CAC96646.1; PID:g16413888; GSPDB:GN00178  
A:Experimental source: strain Cl1p11262  
C:Genetics:  
A:Gene: lisk

	Query March	8.88	Score 87	DB 2	Length 483	
	Best Local Similarity	24.08	Pred. No. 5.7			
	Matches	37	Conservative	29	Mismatches	58
					Indels	30
					Gaps	7
QY	18 KSLQDQDNKRRYYLAQATQTDTGVMGEFTCVSYAEKIGKKLKLATILYKNNKHLTDLKE	77				
	308 KILVQEMLDISRAQISQTKELQITDVNAIVEGVRRNFVEVMEFTFTLKED	359				
QY	78 HETTYAKADYTTENGIKYETQGTQTQTEDYFVFSYDNCVIVYFKERGS	130				
	360 DTDLRALIQNNHLEQIILIIIMDNNAVKYSGGTEVDMHVYKQKQIHL	406				
QY	131 DEGDYELWASEDKTIPDCKFTMAFADQOEK	164				
DB	407 DVRDYGGISQOEIDTKENRFRYVDKASRKEK	438				

RESULT 4  
S18453  
variant surface glycoprotein Mraat 1.6 - Trypanosoma brucei  
C;Species: Trypanosoma brucei

C:Date: 16-Sep-1992 #sequence\_revision 16-Sep-1992 #tex\_change 20-Aug-1999  
C:Accession: S18453  
R:Carrington, M.; Miller, N.; Blum, M.; Roditt, I.; Wiley, D.; Turner, M.  
J. Mol. Biol. 221, 823-835, 1991  
A:Title: Variant specific glycoprotein of Trypanosoma brucei consists of two domains  
A:Reference number: S18445; MUID:92046037; PMID:1942032  
A:Accession: S18453  
A:Molecule type: mRNA  
A:Residues: 1-529 <CAR>  
A:Cross-References: EMBL:X56764; NID:g10474; PIDN:CAA40083.1; PID:g10475  
C:Superfamily: variant surface glycoprotein  
C:Keywords: glycoprotein; membrane protein

Query Match	8.7%	Score 86;	DB 2;	Length 529;
Best Local Similarity	21.4%	Pred. No. 7.8;		
Matches	42;	Conservative	26;	Mismatches 64; Indels 64; Gaps 8.

OY	5	ANEAKLSGYDAMKSLDODONKRRYLACATOTTTDVGWGEFFTCVSVTAEGIKGKKLNATI	64
Db	266	ATRAKTDPELYIWAFAFKNLED-----CESTFTSG-----	294
OY	65	LYKKKHLTDLKESHETITWKKAYDYTTENGILKYEQGRQTGFEDVFVSDYKNODVFEV	124
Db	295	-YSRPSPTTLAADETKTAIK--NIVYQKEKID-QATDK-----DDIKNLDTLF-341	
OY	125	PKRSGSGEGDYELWVSEDKIDKIPDCKF-----TWAYEQAOQEKIVR	167
Db	342	--KKGKDEYPOKLDWAMDMDKDLKDATQINIKKLADIITDRSELNKKVLLYTRBOKEQFLT	399
OY	168	NVTYDSSCKPAPA--ON	182
Db	400	KELKEAQEKATQANON	415

RESULT 5  
T03284  
myoblast c1ty protein - fruit fly (*Drosophila melanogaster*)  
C:Species: *Drosophila melanogaster*  
C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 17-Nov-2000  
C:Accession: T03284  
R:Ericsson, M.R.S.; Galletta, B.J.; Abmayr, S.M.  
J. Cell Biol. 138, 589-603, 1997  
A:Title: Drosophila myoblast c1ty encodes a conserved protein that is essential for m  
A:Reference number: Z14881; MUID:97392689; PMID:9245788  
A:Accession: T03284  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1970 <ER1>  
A:Cross-references: EMBL:AF007805; NID:g2367443; PID:g2367444  
C:Genetics:  
A:Gene: mbc  
A:Cross-references: FlyBase:FBgn0015513  
A:Map position: 3

	Query March	8.6%	Score 85.5	DB 2,	length 1970;	
	Best Local Similarity	21.0%;	Pred.	No. 45;		
	Matches	Conservative	35;	Mismatches	54;	Indels 51; Gaps 6;
QY	44 EFTCVSYAEKIGRKKLNLATILYRNKH-----LTDLKSHETIYVMKAYDYTTENGRTKY	97				
	: : : : : :					
Db	455 EFANIAITMSE---KNVEYSVCVANEQGYLMPGLVSISAGHQIPIDEKXSVYYIHDDRKRM	510				
QY	98 ETGCTRIQTGFEDVFVSDYKNCADVIYFPKKEGSDEGDYELAWSBDKIDKIPDCKCFPMAY	157				
	: : : : : : :					
Db	511 Q-----ETFKIHVPIDEFKCOCHRLFVLKHSNSNE-----QMDRETEK-----PFGLAY	552				
QY	158 FAQQQ-----EKTVRNVYTDSCKCPAAQ	181				
	: : : : : :					
Db	553 VRLMQANGTITTCOGHLLAVYKIDHKRYCDKVANCYLELPATVAELO	599				

RESULT  
AG1085

hypothetical protein lmo0086 [imported] - *Listeria monocytogenes* (strain EGD-e)  
C:Species: *Listeria monocytogenes*  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AG1085  
R:Glaser, P.; Frangoul, L.; Buchrisser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussaugel, O.; Entlan, K.D.; Fsihl, H.; Jones, L.M.; Karel, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Me Ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AG1085  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1959 <GLA>  
A:Cross-references: GB:NC\_003210; PIDN:CAC98301.1; PID:g16409445; GSPDB:GN00177  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: lmo0086

Query Match 8.6%; Score 85; DB 2; Length 1959;  
Best Local Similarity 24.8%; Pred. No. 49;  
Matches 35; Conservative 34; Mismatches 44; Indels 28; Gaps 7;

OY 49 SVTAETIG--KKKLATILYKKNHLDKESHER--ITVWKAYDYTTENGIRKYEQTGTR 103  
DB 1324 AISENICYVDKRVKLVYIGEDKONTYKANKNELFVNVYKADOTITLVNSRVDGLPTR 1363  
OY 104 TQTEDEVFVSFDYKNCDFVYPKERGSDEGDYELWVSEBDIKIPDCKFTWMAVPAQ--Q 161  
DB 1384 DVSTVETSLPDYD---VVKV-----EGKYDLIDGNGEKDKV-----YXSEIK 1423  
OY 162 QEKTVRNV-YTDSCKPAPAQ 181  
DB 1424 TEKSLPSMTSTDYSPVAYGQ 1444

RESULT 7  
706491  
beta-fructofuranosidase (EC 3.2.1.26) - garden pea  
N:Alternate names: Invertase  
C:Species: *Pisum sativum* (garden pea)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 21-May-1999  
C:Accession: T06491  
R:Zheng, L.  
submitted to the EMBL Data Library, January 1996  
A:Reference number: Z15715  
A:Accession: T06491  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-555 <ZHA>  
A:Cross-references: EMBL:X85327; NID:g1160487; PID:g773259  
A:Experimental source: cv. Little Marvel  
C:Genetics:  
A:Note: bfruct  
C:Superfamily: beta-fructofuranosidase  
C:Keywords: glycosidase; hydrolase

Query Match 8.5%; Score 84; DB 2; Length 555;  
Best Local Similarity 23.5%; Pred. No. 12;  
Matches 38; Conservative 22; Mismatches 52; Indels 50; Gaps 8;

OY 2 PTWANEALGSGYODAMKSLQODONKRYVLAQTOTTDGVMGEFFTCVSTAEKIKKLN 61  
DB 174 PTIAKINSSSFRRDPTTS-----WLGR-----DGFWR-----VLIGSKIDTKGM- 212  
OY 62 ATILYKKNHLDLKE-----SHETITWKAYDY-----TTENG---IKYE 98  
DB 213 -AIIYKSNFVDWYAKHPLHSABEGTGMWECDFPVLDKMLRTGVDTSRKGGDDVRRHV 271  
OY 99 TQGTTRTQTEDEVFVSFDYKNCDFVYPKERGSDEG-----DY 135

DB 272 LKVSIDTKHHDHYLIGSYDVVKDYFVPENGEDNGFVLRYDY 313

## RESULT 8

T06380  
beta-fructofuranosidase (EC 3.2.1.26) - garden pea  
N:Alternate names: cell wall invertase  
C:Species: *Pisum sativum* (garden pea)  
C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 22-Jun-1999  
C:Accession: T06380  
R:Zheng, L.; Conn, N.S.; Mitchell, J.P.  
submitted to the EMBL Data Library, May 1998  
A:Reference number: Z15639  
A:Accession: T06380  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-582 <ZHA>  
A:Cross-references: EMBL:AF063246; NID:g3152879; PIDN:AACT17166.1; PID:g3152880  
A:Experimental source: cv. Little Marvel  
C:Genetics:  
A:Gene: bfruct  
A:Function: hydrolyzes sucrose to glucose and fructose  
C:Superfamily: beta-fructofuranosidase  
C:Keywords: glycosidase; hydrolase

Query Match 8.5%; Score 84; DB 2; Length 582;  
Best Local Similarity 23.5%; Pred. No. 13;  
Matches 38; Conservative 22; Mismatches 52; Indels 50; Gaps 8;

OY 2 PTWANEALGSGYODAMKSLQODONKRYVLAQTOTTDGVMGEFFTCVSTAEKIKKLN 61  
DB 174 PTIAKINSSSFRRDPTTS-----WLGR-----DGFWR-----VLIGSKIDTKGM- 212  
OY 62 ATILYKKNHLDLKE-----SHETITWKAYDY-----TTENG---IKYE 98  
DB 213 -AIIYKSNFVDWYAKHPLHSABEGTGMWECDFPVLDKMLRTGVDTSRKGGDDVRRHV 271  
OY 99 TQGTTRTQTEDEVFVSFDYKNCDFVYPKERGSDEG-----DY 135  
DB 272 LKVSIDTKHHDHYLIGSYDVVKDYFVPENGEDNGFVLRYDY 313

## RESULT 9

T00494  
RNA-directed RNA polymerase (EC 2.7.7.48) - white button mushroom virus 1  
N:Alternate names: RNA-dependent RNA polymerase  
C:Species: white button mushroom virus 1  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 08-Oct-1999  
C:Accession: T00494  
R:van der Lende, T.R.; Duitman, E.H.; Gunnwylk, M.G.W.; Yu, L.; Wessels, J.G.H.  
Virology 217, 88-96, 1996  
A:Title: Functional analysis of dsRNAs (L1, L3, L5 and M2) associated with isometric 3  
A:Reference number: Z14158; MUID:96177126; PMID:8599239  
A:Accession: T00494  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: genomic RNA  
A:Residues: 1-1078 <VAN>  
A:Cross-references: EMBL:X94361; NID:g1217917; PIDN:CAAG4144.1; PID:g219181  
A:Genetics:  
A:Gene: L1  
C:Keywords: nucleotidyltransferase

Query Match 8.4%; Score 83.5; DB 2; Length 1078;  
Best Local Similarity 21.4%; Pred. No. 32;  
Matches 39; Conservative 33; Mismatches 71; Indels 39; Gaps 9;

OY 16 AAKSLQODONKRYVLAQTOTTDGVMGEFFTCVSTAEKIKKLNATILYKN----- 68  
DB 53 ALSNINPDNAPFIVQICELAEAL-GVQIANVLLVKEGI-----IQGIDRYKSLSRECK 107  
OY 69 --KHLTDKESHERITWKAYDY-----TTENGIRYEQGTQTEDEVFVSFDYKNCDFV 121

Db 108 TLKHVHVNHNHNLWKEK-YKTHRVVLSYNDIYHKVGLIYISV---GCYKSSFGEDDV 163  
QY 122 IFVFKERGSDEGDEYELWSEDKIDKIPD-----CCKFTMAVFAOQDEKTVRNVTDS 173  
Db 164 L-----SGIDERTYD---AREKVNESNLKESKHPATKAVCAKYLQONFOAAPADYANA 215  
QY 174 SC 175  
Db 216 SC 217

RESULT 10  
112821  
hypothetical protein yong - Bacillus subtilis phage SPBC2  
C:Species: Bacillus subtilis phage SPBC2  
C>Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 15-Oct-1999  
C:Accession: T12821; A69914  
R:Lazarovic, V.; Duesterhoeft, A.; Soldo, B.; Hilbert, H.; Manuel, C.; Karamata, D.  
submitted to the EMBL Data Library, August 1997  
A:Description: The complete nucleotide sequence of the Bacillus subtilis SPbetac2 prophage  
A:Reference number: 217583  
A:Accession: T12821  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-306 <LNA>  
A:Cross-references: EMBL:AF020713; NID:g3023478; PID:g3025535; PIDN:AC13030.1  
R:Kunst, F.; Ogawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter  
C.; Bron, S.; Brouillet, S.; Brusch, L.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chc  
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fultz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallazzi, A.; Galler  
lech, J.; Harwood, C.R.; Henauf, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullo, M.F.  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A:Authors: Labber, J.; Lazarovic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel  
Y, M.; Ogawa, K.; Ogawara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli  
Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A:Reference number: A69580; MUID:98044033; PMID:9384377  
A:Accession: A69914  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-306 <KNC>  
A:Cross-references: GB:299115; GB:AL009126; NID:g2634478; PIDN:CAB14028.1; PID:e1183557;  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yong

Query Match 8.3%; Score 82.5; DB 2; Length 306;  
Best Local Similarity 22.8%; Pred. No. 8;  
Matches 37; Conservative 20; Mismatches 56; Indels 49; Gaps 6;

QY 48 VSTAEIKGKKLNATL---YKTKHTDLKESHETTVKAKADYTTENGIKKFTGTRT 104  
Db 72 VSAVNESESKKSGAGNLTGYMKNGKDYKSK---TWADSEFDEEEDQEVYTTQLLA 126  
QY 105 QTEFEDVVFSDYKNCVDYIFVFKERGSDEGDEYELWSEDKIDKIPDCKFTMAV----- 157  
Db 127 KSTEDI---SKEDIDELMPGKGLDVEDY-TWLQNEYID-----FTNRRECDSKGM 174  
QY 158 -----FAOQDEKTVRNVTDSCKP 177  
Db 175 ELLINEICLTRLDIRKRENGEKVDDQKTLQDLGLSSNLRP 216

RESULT 11  
G71616  
hypothetical protein PFB0375w - malaria parasite (Plasmodium falciparum)  
C:Species: Plasmodium falciparum  
C>Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 21-Jul-2000

C:Accession: G71616  
R:Gardner, M.J.; Tetteh, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.  
A:Parta, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H  
Science 282, 1126-1132, 1998  
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.  
A:Reference number: A71600; MUID:99021743; PMID:9804551  
A:Accession: G71616  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1802 <GAR>  
A:Cross-references: GB:AE001390; GB:AE001362; NID:g3845164; PIDN:AC71864.1; PID:g384  
A:Experimental source: clone 307  
C:Genetics:  
A:Gene: PFB0375w

Query Match 8.3%; Score 82.5; DB 2; Length 1802;  
Best Local Similarity 25.9%; Pred. No. 74;  
Matches 30; Conservative 13; Mismatches 42; Indels 31; Gaps 4;

QY 61 NATLYKKNHLTDKESHETTVKAYDYTTENGIKKFTGTRT-----RTQTEFEDVVFSDY 116  
Db 21 NNNITGSDYDPLKELTYTNHYKSYHKKENIKYKAEKENVIDINNNKDDLFYNDH 80  
QY 117 KNCVDYIFVFKERGSDEGDEYELWSEDKIDKIPDCKFTMAVFAOQDEKTVRNVTDS 172  
Db 81 K-----IDDKIRKRI-QCKNCTTY-----HDIEKNSYVD 109

RESULT 12  
146001  
C4b-binding protein alpha chain - bovine  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: I46001; S43190  
R:Hillarp, A.; Thern, A.; Dahlback, B.  
J. Immunol. 153, 4190-4199, 1994  
A:Title: Bovine C4b binding protein. Molecular cloning of the alpha- and beta-chains  
A:Reference number: I46001; MUID:95015909; PMID:7930621  
A:Accession: I46001  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-610 <HL>  
A:Cross-references: EMBL:Z31693; NID:g469117; PIDN:CAA83498.1; PID:g469118  
C:Superfamily: C4b-binding protein alpha chain; complement factor H repeat homology  
F:50-107/Domain: complement factor H repeat homology <FH>  
F:112-169/Domain: complement factor H repeat homology <FH2>  
F:174-234/Domain: complement factor H repeat homology <FH3>  
F:239-284/Domain: complement factor H repeat homology <FH4>  
F:299-362/Domain: complement factor H repeat homology <FH5>  
F:366-425/Domain: complement factor H repeat homology <FH6>  
F:429-483/Domain: complement factor H repeat homology <FH7>  
F:487-541/Domain: complement factor H repeat homology <FH8>

Query Match 8.2%; Score 81; DB 1; Length 610;  
Best Local Similarity 23.7%; Pred. No. 26;  
Matches 32; Conservative 14; Mismatches 47; Indels 42; Gaps 5;

QY 78 HETTWKAYD-----YTTENGIKKFTGTRTQTEFEDVVFSDYKNCVDYIFVFKERGS 131  
Db 315 HQOTEBQVYDIDGFLVSKHFKGKRETDPTVTQSNLEMSPIECKKCCPERPLNN 374  
QY 132 EGDYEL-----WSEDKID-----KIPDC---CKFTMA 156  
Db 375 YGSITTLRRRSTSTHCHYISGDKISVESHCKYMFDAICTKHGTWSPRTPECRDCK-SP 433  
QY 157 YFAOQDEKTVRNVT 171  
Db 434 VIAHGKRVSKFT 448

RESULT 13  
S30484  
pol polyprotein - human immunodeficiency virus type 2

C:Species: human immunodeficiency virus type 2, HIV-2  
 C>Date: 02-Dec-1993 #sequence\_revision 01-Dec-1995 #text\_change 23-Mar-2001  
 C:Accession: S30484  
 R:Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;  
 submitted to the EMBL Data Library, December 1992  
 A:Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa  
 A:Reference number: S30460  
 A:Accession: S30484  
 A:Status: preliminary  
 A:Molecule type: nucleic acid  
 A:Residues: 1-656 <GAG>  
 A:Cross-references: EMBL:M87114  
 C:Superfamily: pol polyprotein

Query Match 8.2%; Score 81; DB 2; Length 656;  
 Best Local Similarity 23.0%; Pred. NO. 28; Mismatches 50; Indels 66; Gaps 9;

Matches 42; Conservative 25; Mismatches 50; Indels 66; Gaps 9;  
 Oy 15 DAMSLAODQNKRYVLAQATQTTDGVWGEFTCVSTAETKIGKKLNATILYKKNHLDL 74  
 Db 126 DAFSIPDEEPRQTA-----FTLPSVNNAEPRGKR-----YLYK-----VL 162  
 Oy 75 KESHEITVWKA-----YTTTENG-----KETQGTQTQTFEDVFEVDYKNC---YIF 123  
 Db 163 PQG-----WKGSPALFQYIMRNVLPEPRKANSVDTLIQYMDLILASDRDLHDHRYVL 216  
 Oy 124 VPKRGSDG-----DYELWVSEDKIDKIPDCKFMAYFAQOQEKTV 166  
 Db 217 QLKELNDLGLFTEPEKQKPRPYHMWGYELMPTKMKLQKTE-----LPQKEDWTV 267  
 Oy 167 RNW 169  
 Db 268 NDI 270

# RESULT 14

S73939  
 oligopeptide transport ATP-binding protein oppF - Mycoplasma pneumoniae (strain ATCC 293  
 N:Alternate names: hypothetical protein G07\_orf81  
 C:Species: Mycoplasma pneumoniae  
 A:Date: 27-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 02-Feb-2001  
 C:Accession: S73939  
 R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Plöckl, E.; Li, B.C.; Herrmann, R.  
 Nucleic Acids Res. 24, 4420-4449, 1996  
 A>Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae  
 A:Reference number: S73327; MUID:97105885; PMID:8948633  
 A:Accession: S73939  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-851 <HIM>  
 A:Cross-references: EMBL:AE000058; GB:U00089; NID:q1674291; PIDN:AAB96261.1; PID:q167431  
 A>Note: The nucleotide sequence was submitted to the EMBL Data Library, November 1996  
 C:Genetics:  
 A:Gene: oppF  
 A:Genetic code: SGC3  
 C:Superfamily: unassigned ATP-binding cassette proteins: ATP-binding cassette homology  
 C:Keywords: ATP; nucleotide binding; P-loop  
 F:31764/Domain: ATP-binding cassette homology #status atypical <ABCL>  
 F:48-55/Region: nucleotide-binding motif A (P-loop)

Query Match 8.2%; Score 81; DB 2; Length 851;  
 Best Local Similarity 23.7%; Pred. NO. 39;

Matches 44; Conservative 26; Mismatches 60; Indels 56; Gaps 8;

Oy 3 TWANFAKLSYODA-WKSLQ-DQNKRYVLAQATQTTDGVWGEFTCVSTAETKIGKK- 59  
 Db 251 TTOOKAKYSTOAFATWKLQTKONLKAIRAO-----MAEELQNKPR 292  
 Oy 60 --LNAITLYKKNHLDLAKESHETIVWKAIDYTTENGIKYETQGTQTQTFEDVFEVS--- 114  
 Db 293 IYLNAML-----TTKNYIKDSRQ--NTQLTDVFAFSYND 326

Oy 115 --DYKNCDFVFPKRGSDGDEYELWSE--DKIDKIPDCKFTMAVFAQOQEKTVRN 170  
 Db 327 MVDKRRRLVVLSEYKALPYFYDWINQNDPRDELINAFPLDIDVYIALNDFANVE 386  
 Oy 171 TDSSCK 176  
 Db 387 SDAKAE 392

# RESULT 15

138614  
 helicase II - human  
 C:Species: Homo sapiens (man)  
 C>Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 05-Nov-1999  
 C:Accession: 138614  
 R:Stayton, C.L.; Dabovic, B.; Gullano, M.; Geetz, J.; Broccoli, V.; Giovannazzi, S.; B  
 Hum. Mol. Genet. 3, 1957-1964, 1994  
 A>Title: Cloning and characterization of a new human Xq13 gene, encoding a putative h  
 A:Reference number: 138614; MUID:95179111; PMID:7874112  
 A:Accession: 138614  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1641 <RES>  
 A:Cross-references: EMBL:U09820; NID:q606832; PIDN:AAC50069.1; PID:q606833  
 C:Genetics:  
 A:Gene: RAD54L

Query Match 8.2%; Score 81; DB 2; Length 1641;  
 Best Local Similarity 26.4%; Pred. NO. 89;

Matches 42; Conservative 21; Mismatches 56; Indels 40; Gaps 8;

Oy 32 QATQTTDGV---WGEFTCVSTAETKIGKKLNATILYKKNHLDLAKESHETIVWKA 88  
 Db 54 QASASTDGVDLKSKRGQSFSLSEVKKVAETK-----EKSRL-----KTCTCKKQD 100  
 Oy 89 YTTENGIRYETQGTQTFEDVFEVSDYKNCDFVFPKRGSD-----EGDYELWS 140  
 Db 101 GLSDIAERFLKKGDSDETSED-----DKQSKGTEKKKPSDFKKVYKMQYE--SS 153  
 Oy 141 EDKIDKIP---DCKFTMAVFAQOQEKTVRNVTYDSSCK 176  
 Db 154 SDGTEKLPERREICHFPKGI-----KQIKNGTTDGERK 186

Search completed: August 1, 2003, 12:16:36  
 Job time : 15.52 secs

**THIS PAGE BLANK (USPTO)**